SCORE OVER LENGTH SEARCHES

Attached is a score over length search. This search was developed to overcome limitations in most standard search systems which favor large sequences with high scoring, but lesser overall identity over smaller sequences with higher overall identity. This search is especially useful for relatively small nucleic acid or polypeptide target sequences (antisense, fragments, probes, primers, RNAi, epitopes, haptens, etc.) claimed functionally via a form of hybridization and/or identity language and having defined upper and lower polynucleotide and or polypeptide length limits.

The score over length search is performed by first running the query sequence using examiner-specified identity and polynucleotide or protein length limit parameters, and saving 65,000 hits and 0 alignments from each desired database. The resulting output is reformatted using a Microsoft Word macro and is imported into Excel. The summary table data are then sorted by the ratio of score of each hit sequence divided by its length and the accession numbers for all hits below the examiner's desired score over length parameters are deleted. The remaining accession numbers are used to pull the corresponding sequences from the databases into subdatabases enriched for good hits and the query sequence is re-run against these subdatabases to yield the final results.

The score over length cutoff for this search is 80%.

Examiner Please Note: This cover sheet should be included when submitting results to be scanned.

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 26, 2005, 14:59:18 ; Search time 3132 Seconds
(without alignments)
230.914 Million cell updates/sec

Title: US-10-619-906-5
Perfect score: 19
Sequence: 1 aattattgattctgtaggt 19

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		19	100.0	208	2	BF327490	BF327490 QV4-BN009
2		19	100.0	255	2	BF088326	BF088326 CM1-HT087
3		19	100.0	255	2	BF155525	BF155525 CM1-HT087
4	c	19	100.0	385	2	AW996471	AW996471 QV3-BN004
5	c	19	100.0	442	2	AW996697	AW996697 QV3-BN004
6	c	19	100.0	481	5	BQ312455	BQ312455 MR0-BN011
7	c	19	100.0	568	2	BE440047	BE440047 HTM1-818F
8	c	19	100.0	583	5	BP352329	BP352329 BP352329
9	c	19	100.0	664	2	BF327500	BF327500 QV4-BN009
10	c	19	100.0	714	1	AV716402	AV716402 AV716402
11	c	19	100.0	930	5	BX377010	BX377010 BX377010
12	c	19	100.0	1784	3	CR603756	CR603756 full-leng
13	c	17	89.5	495	8	AQ278111	AQ278111 CITBI-E1-
14	c	17	89.5	528	8	BH886093	BH886093 LB00584a.
15	c	17	89.5	631	8	BH888327	BH888327 LB01983a.
16	c	17	89.5	645	8	BH885870	BH885870 LB00450a.
17	c	17	89.5	808	7	CR575999	CR575999 CR575999
18	c	17	89.5	885	5	BX710423	BX710423 BX710423
19		16	84.2	251	6	CA049038	CA049038 ssalkspa0
20		16	84.2	273	2	BB127347	BB127347 BB127347
21		16	84.2	298	1	AV296455	AV296455 AV296455
22	c	16	84.2	312	2	BB040926	BB040926 BB040926
23	c	16	84.2	375	7	CV317979	CV317979 CM2-GN005
24		16	84.2	410	4	BG360551	BG360551 RTI00001
25		16	84.2	457	1	AI760729	AI760729 wi08f05.x
26	c	16	84.2	506	6	CA061810	CA061810 ssalrgp00
27	c	16	84.2	562	7	CO228348	CO228348 WS0014.B2
28		16	84.2	566	7	CK881762	CK881762 SGP142112
29	c	16	84.2	581	2	AW912973	AW912973 uf47a10.y
30		16	84.2	581	4	BI946345	BI946345 01088 lea
31		16	84.2	583	9	CR315576	CR315576 Medicago
32	c	16	84.2	591	6	CA043357	CA043357 ssalplnb5
33		16	84.2	611	6	CD334984	CD334984 StrPu536.
34		16	84.2	622	6	CA353561	CA353561 625083 NC
35	c	16	84.2	634	7	CO217921	CO217921 WS0101.B2
36		16	84.2	642	6	CB502289	CB502289 ssalplnb5
37		16	84.2	647	9	AG151770	AG151770 Pan trogl
38		16	84.2	648	6	CA377750	CA377750 656393 NC
39		16	84.2	651	5	BX077071	BX077071 BX077071
40	c	16	84.2	651	9	CE160496	CE160496 tigr-gss-
41	c	16	84.2	655	6	CB497718	CB497718 omykrbna5
42	c	16	84.2	669	9	CE545000	CE545000 tigr-gss-
43	c	16	84.2	683	7	CO204386	CO204386 WS0064.B2
44	c	16	84.2	686	8	BH973406	BH973406 odj04e11.
45	c	16	84.2	687	7	CO240857	CO240857 WS0078.B2

ALIGNMENTS

RESULT 1
BF327490
LOCUS BF327490 208 bp mRNA linear EST 22-NOV-2000
DEFINITION QV4-BN0090-210400-182-a11 BN0090 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF327490
VERSION BF327490.1 GI:11298238
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 208)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV4&t2=QV4-BN0090-210400-182-a11&t3=2000-04-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 4.
FEATURES Location/Qualifiers
source 1. .208
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0090"
/note="Organ: breast_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        157 AATTATTGATTCTGTAGGT 175
```

RESULT 2
BF088326
LOCUS BF088326 255 bp mRNA linear EST 19-OCT-2000
DEFINITION CM1-HT0876-060900-392-h07 HT0876 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF088326
VERSION BF088326.1 GI:10894036
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 255)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM1-HT0876-060900-392-h07&t3=2000-09-06&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 255.
FEATURES Location/Qualifiers
source 1. .255
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0876"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db         46 AATTATTGATTCTGTAGGT 64
```

RESULT 3
BF155525
LOCUS BF155525 255 bp mRNA linear EST 30-OCT-2000
DEFINITION CM1-HT0876-260900-392-h07 HT0876 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF155525
VERSION BF155525.1 GI:11050708
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 255)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM1-HT0876-260900-392-h07&t3=2000-09-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 255.
FEATURES Location/Qualifiers
source 1. .255
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0876"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db         46 AATTATTGATTCTGTAGGT 64
```

RESULT 4
AW996471/c
LOCUS AW996471 385 bp mRNA linear EST 05-JUN-2000
DEFINITION QV3-BN0046-150300-122-c09 BN0046 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW996471

```
VERSION     AW996471.1  GI:8256705
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 385)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV3-BN0046-150
            300-122-c09&t3=2000-03-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 385.
FEATURES             Location/Qualifiers
     source          1. .385
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="BN0046"
                     /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 385;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        303 AATTATTGATTCTGTAGGT 285


RESULT 5
AW996697/c
LOCUS       AW996697                442 bp    mRNA    linear   EST 05-JUN-2000
DEFINITION  QV3-BN0046-150400-151-h01 BN0046 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW996697
VERSION     AW996697.1  GI:8256931
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV3-BN0046-150
            400-151-h01&t3=2000-04-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 7
            High quality sequence stop: 442.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="BN0046"
                     /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 442;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        244 AATTATTGATTCTGTAGGT 226


RESULT 6
BQ312455/c
LOCUS       BQ312455                481 bp    mRNA    linear   EST 16-MAY-2002
DEFINITION  MR0-BN0115-041000-014-f08 BN0115 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ312455
VERSION     BQ312455.1  GI:20868627
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
```

```
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR0&t2=MR0-BN0115-
            041000-014-f08&t3=2000-10-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 18
            High quality sequence stop: 47.
FEATURES              Location/Qualifiers
     source           1. .481
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="BN0115"
                      /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                      Site_2: SmaI; A mini-library was made by cloning products
                      derived from  ORESTES PCR (U.S. Letters Patent application
                      No. 196,716 -  Ludwig Institute for Cancer Research)
                      profiles into  the pUC 18 vector. Reverse transcription of
                      tissue mRNA and cDNA amplification were performed under
                      low stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 5;  Length 481;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        382 AATTATTGATTCTGTAGGT 364


RESULT 7
BE440047/c
LOCUS       BE440047                568 bp    mRNA    linear   EST 25-JUL-2000
DEFINITION  HTM1-818F HTM1 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE440047
VERSION     BE440047.1  GI:9439531
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Gonzalez,P., Epstein,D.L. and Borras,T.
  TITLE     Characterization of gene expression in human trabecular meshwork
            using single-pass sequencing of 1060 clones
  JOURNAL   Invest. Ophthalmol. Vis. Sci. (2000) In press
COMMENT     Contact: Pedro Gonzalez
            Department of Ophthalmology
            Duke University
            Duke Eye Center, Erwin Rd, Box 3802, Durham, NC 27710-3802, USA
            Tel: 919 681 4085
            Fax: 919 684 8983
            Email: pedro.gonzalez@duke.edu.
FEATURES              Location/Qualifiers
     source           1. .568
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /tissue_type="Trabecular meshwork"
                      /clone_lib="HTM1"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 568;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        408 AATTATTGATTCTGTAGGT 390


RESULT 8
BP352329/c
LOCUS       BP352329                583 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP352329 Sugano cDNA library, squamous cell TE13 Homo sapiens cDNA
            clone T3R07476, mRNA sequence.
ACCESSION   BP352329
VERSION     BP352329.1  GI:52282315
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES              Location/Qualifiers
     source           1. .583
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="T3R07476"
                      /cell_type="squamous cell"
                      /cell_line="TE13"
                      /clone_lib="Sugano cDNA library, squamous cell TE13"
                      /note="well-differentiated squamous cell carcinoma"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 5;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        405 AATTATTGATTCTGTAGGT 387


RESULT 9
BF327500/c
LOCUS       BF327500                664 bp    mRNA    linear   EST 22-NOV-2000
DEFINITION  QV4-BN0090-220700-307-c09 BN0090 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF327500
VERSION     BF327500.1  GI:11298248
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
```

```
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV4&t2=QV4-BN0090-
            220700-307-c09&t3=2000-07-22&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="BN0090"
                     /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 664;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        423 AATTATTGATTCTGTAGGT 405


RESULT 10
AV716402/c
LOCUS       AV716402                  714 bp    mRNA    linear   EST 11-OCT-2000
DEFINITION  AV716402 DCB Homo sapiens cDNA clone DCBAQG01 5', mRNA sequence.
ACCESSION   AV716402
VERSION     AV716402.1  GI:10797919
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 714)
  AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
  TITLE     Homo sapiens cDNA DCB clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .714
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBAQG01"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 1;  Length 714;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        352 AATTATTGATTCTGTAGGT 334


RESULT 11
BX377010/c
LOCUS       BX377010                  930 bp    mRNA    linear   EST 26-APR-2004
DEFINITION  BX377010 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI001YG13 5-PRIME, mRNA sequence.
ACCESSION   BX377010
VERSION     BX377010.2  GI:46570565
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30442870.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            5783.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI001AD07QP1&c=5783.r.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI001YG13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 5;  Length 930;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches.  19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        396 AATTATTGATTCTGTAGGT 378


RESULT 12
CR603756/c
LOCUS       CR603756                 1784 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI002YD16 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR603756
VERSION     CR603756.1  GI:50484563
KEYWORDS    HTC; CNSLT_cDNA.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1784)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1784)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1784
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI002YD16"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 3;  Length 1784;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        394 AATTATTGATTCTGTAGGT 376


RESULT 13
AQ278111/c
LOCUS       AQ278111                495 bp    DNA     linear   GSS 22-NOV-1998
DEFINITION  CITBI-E1-2523H15.TR CITBI-E1 Homo sapiens genomic clone 2523H15,
            genomic survey sequence.
ACCESSION   AQ278111
VERSION     AQ278111.1  GI:3904154
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CITBI-E1-2523H15.TF
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2523H15"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 8;  Length 495;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db        230 TTATTGATTCTGTAGGT 214


RESULT 14
BH886093/c
LOCUS       BH886093                528 bp    DNA     linear   GSS 07-AUG-2002
DEFINITION  LB00584a.d_SP6.1 Leishmania major Friedlin BAC Library Leishmania
            major genomic clone LB00584a, genomic survey sequence.
ACCESSION   BH886093
VERSION     BH886093.1  GI:22130488
KEYWORDS    GSS.
SOURCE      Leishmania major
  ORGANISM  Leishmania major
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Leishmania.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Myler,P.J., Vogt,C., Munden,H., Robertson,L., Sisk,E.,
            Fazelinia,G., Aggarwal,G., Nelson,S., Seyler,A., Worthey,E.,
            Stuart,K. and Ragland,M.
  TITLE     Leishmania major Friedlin BAC End Sequences
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: LB00584a.d_T7.1
            Contact: Myler PJ
            Seattle Biomedical Research Institute
            4 Nickerson Street, Seatttle, WA 98109-1651, USA
            Tel: 206 284-8846
            Fax: 206 284-0313
            Email: mylerpj@sbri.org
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Leishmania major"
                     /mol_type="genomic DNA"
                     /strain="Friedlin"
                     /db_xref="taxon:5664"
                     /clone="LB00584a"
                     /lab_host="E. coli GeneHogs + TrfA"
                     /clone_lib="Leishmania major Friedlin BAC Library"
                     /note="Vector: pCG270; Site_1: HindIII; Genomic DNA from
                     Leishmania major Friedlin in agarose blocks was partially
                     digested with HindIII, size selected, and ligated with
                     HindIII-digested pCG270 vector DNA. 10368 clones were
                     picked and arrayed in 384- and 96-will plates. Library
                     construction and arraying was carried out by ResGen
                     Corporation and clones and filters are available from
                     them"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 8;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
```

```
                  ||||||||||||||||||
Db        260 TTATTGATTCTGTAGGT 244


RESULT 15
BH888327/c
LOCUS       BH888327                631 bp    DNA     linear   GSS 07-AUG-2002
DEFINITION  LB01983a.d_T7.1 Leishmania major Friedlin BAC Library Leishmania
            major genomic clone LB01983a, genomic survey sequence.
ACCESSION   BH888327
VERSION     BH888327.1  GI:22134207
KEYWORDS    GSS.
SOURCE      Leishmania major
  ORGANISM  Leishmania major
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Leishmania.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Myler,P.J., Vogt,C., Munden,H., Robertson,L., Sisk,E.,
            Fazelinia,G., Aggarwal,G., Nelson,S., Seyler,A., Worthey,E.,
            Stuart,K. and Ragland,M.
  TITLE     Leishmania major Friedlin BAC End Sequences
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: LB01983a.d_SP6.1
            Contact: Myler PJ
            Seattle Biomedical Research Institute
            4 Nickerson Street, Seatttle, WA 98109-1651, USA
            Tel: 206 284-8846
            Fax: 206 284-0313
            Email: mylerpj@sbri.org
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Leishmania major"
                     /mol_type="genomic DNA"
                     /strain="Friedlin"
                     /db_xref="taxon:5664"
                     /clone="LB01983a"
                     /lab_host="E. coli GeneHogs + TrfA"
                     /clone_lib="Leishmania major Friedlin BAC Library"
                     /note="Vector: pCG270; Site_1: HindIII; Genomic DNA from
                     Leishmania major Friedlin in agarose blocks was partially
                     digested with HindIII, size selected, and ligated with
                     HindIII-digested pCG270 vector DNA. 10368 clones were
                     picked and arrayed in 384- and 96-will plates. Library
                     construction and arraying was carried out by ResGen
                     Corporation and clones and filters are available from
                     them"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 8;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db        256 TTATTGATTCTGTAGGT 240


RESULT 16
BH885870/c
LOCUS       BH885870                645 bp    DNA     linear   GSS 07-AUG-2002
DEFINITION  LB00450a.d_T7.1 Leishmania major Friedlin BAC Library Leishmania
            major genomic clone LB00450a, genomic survey sequence.
ACCESSION   BH885870
VERSION     BH885870.1  GI:22130265
KEYWORDS    GSS.
SOURCE      Leishmania major
  ORGANISM  Leishmania major
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Leishmania.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Myler,P.J., Vogt,C., Munden,H., Robertson,L., Sisk,E.,
            Fazelinia,G., Aggarwal,G., Nelson,S., Seyler,A., Worthey,E.,
            Stuart,K. and Ragland,M.
  TITLE     Leishmania major Friedlin BAC End Sequences
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: LB00450a.d_SP6.1
            Contact: Myler PJ
            Seattle Biomedical Research Institute
            4 Nickerson Street, Seatttle, WA 98109-1651, USA
            Tel: 206 284-8846
            Fax: 206 284-0313
            Email: mylerpj@sbri.org
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Leishmania major"
                     /mol_type="genomic DNA"
                     /strain="Friedlin"
                     /db_xref="taxon:5664"
                     /clone="LB00450a"
                     /lab_host="E. coli GeneHogs + TrfA"
                     /clone_lib="Leishmania major Friedlin BAC Library"
                     /note="Vector: pCG270; Site_1: HindIII; Genomic DNA from
                     Leishmania major Friedlin in agarose blocks was partially
                     digested with HindIII, size selected, and ligated with
                     HindIII-digested pCG270 vector DNA. 10368 clones were
                     picked and arrayed in 384- and 96-will plates. Library
                     construction and arraying was carried out by ResGen
                     Corporation and clones and filters are available from
                     them"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 8;  Length 645;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db        256 TTATTGATTCTGTAGGT 240


RESULT 17
CR575999/c
LOCUS       CR575999                808 bp    mRNA    linear   EST 21-JUL-2004
DEFINITION  CR575999 XGC-tailbud-head Xenopus tropicalis cDNA clone THdA040j06
            5', mRNA sequence.
ACCESSION   CR575999
VERSION     CR575999.1  GI:50463425
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 808)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (2004)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Croning MDR
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
             TROPICALIS_SEQUENCE_ID: THdA040j06.p1kaSP6
             This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Nigel Garrett.
            Seq primer: SP6.
FEATURES             Location/Qualifiers
     source          1. .808
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="THdA040j06"
```

/dev_stage="tailbud head (stage 28-30)"
/lab_host="Escherichia coli DH10B."
/clone_lib="XGC-tailbud-head"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dT primed from 5ug of poly A+ RNA from tailbud head. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 89.5%; Score 17; DB 7; Length 808;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 ATTATTGATTCTGTAGG 18
              |||||||||||||||||
Db        597 ATTATTGATTCTGTAGG 581
```

RESULT 18
BX710423/c
LOCUS BX710423 885 bp mRNA linear EST 18-NOV-2003
DEFINITION BX710423 XGC-tadpole Xenopus tropicalis cDNA clone TTpA003j14 5', mRNA sequence.
ACCESSION BX710423
VERSION BX710423.1 GI:38382526
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 885)
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TTpA003j14.p1kSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Nigel Garrett.
cDNA was oligo dT primed from 5ug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B.
FEATURES Location/Qualifiers
source 1. .885
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTpA003j14"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dT primed from 5ug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 89.5%; Score 17; DB 5; Length 885;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 ATTATTGATTCTGTAGG 18
              |||||||||||||||||
Db        740 ATTATTGATTCTGTAGG 724
```

RESULT 19
CA049038
LOCUS CA049038 251 bp mRNA linear EST 04-MAR-2003
DEFINITION ssalkspa003036 kidney Salmo salar cDNA, mRNA sequence.
ACCESSION CA049038
VERSION CA049038.1 GI:24355208
KEYWORDS EST.
SOURCE Salmo salar (Atlantic salmon)
ORGANISM Salmo salar
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE 1 (bases 1 to 251)
AUTHORS GRASP Consortium, Davidson,W.S., Koop,B.F. and http://web.uvic.ca/cbr/grasp.
TITLE A survey of Salmo salar transcripts from high complexity cDNA libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@uvic.ca
Centre for Biomedical Research, University of Victoria
cDNA preparation and sequencing:
Roberto Alberto, Marianne Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.
bioinformatics:
Gordon D Brown.
FEATURES Location/Qualifiers
source 1. .251
/organism="Salmo salar"
/mol_type="mRNA"
/strain="McConnell"
/db_xref="taxon:8030"
/clone_lib="kidney"
/note="Vector: pBlueScriptIISK+; Library Creator: Matthew L Rise ; Atlantic salmon tissue contributors: Carlo Biagi, Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.), Simon Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery (Crofton, B.C.), Rachel Roper (University of Victoria)"

ORIGIN

Query Match 84.2%; Score 16; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        113 TATTGATTCTGTAGGT 128
```

RESULT 20
BB127347
LOCUS BB127347 273 bp mRNA linear EST 28-JUN-2000
DEFINITION BB127347 RIKEN full-length enriched, 16 days neonate cerebellum Mus musculus cDNA clone 9630021A03 3', mRNA sequence.
ACCESSION BB127347
VERSION BB127347.1 GI:8781701
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 273)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,

Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Konno,H., et al.)

JOURNAL Unpublished (2000)

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .273
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9630021A03"
/tissue_type="cerebellum"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 16 days neonate cerebellum"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 84.2%; Score 16; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTA 16
              ||||||||||||||||
Db        173 AATTATTGATTCTGTA 188
```

RESULT 21
AV296455

LOCUS AV296455 298 bp mRNA linear EST 10-NOV-1999

DEFINITION AV296455 RIKEN full-length enriched, 8 days embryo Mus musculus cDNA clone 5730441N12 3', mRNA sequence.

ACCESSION AV296455

VERSION AV296455.1 GI:6328474

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 298)

AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Konno,H., et al. 1999)

JOURNAL Unpublished (1999)

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .298
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5730441N12"
/sex="mixed"
/dev_stage="8 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 8 days embryo"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 100.0 Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda

FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTA 16
||||||||||||||||
Db 196 AATTATTGATTCTGTA 211

RESULT 22
BB040926/c
LOCUS BB040926 312 bp mRNA linear EST 23-JUN-2000
DEFINITION BB040926 RIKEN full-length enriched, 13 days embryo male testis Mus musculus cDNA clone 6030452I03 3', mRNA sequence.
ACCESSION BB040926
VERSION BB040926.1 GI:8447312
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 312)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .312
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="6030452I03"
/sex="male"
/tissue_type="testis"
/dev_stage="13 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 13 days embryo male testis"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
ORIGIN

Query Match 84.2%; Score 16; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTA 16
||||||||||||||||
Db 39 AATTATTGATTCTGTA 24

RESULT 23
CV317979/c
LOCUS CV317979 375 bp mRNA linear EST 24-SEP-2004
DEFINITION CM2-GN0055-060900-356-e10 GN0055 Homo sapiens cDNA, mRNA sequence.
ACCESSION CV317979
VERSION CV317979.1 GI:52641193
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 375)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. http://www.ludwig.org.br.
FEATURES Location/Qualifiers
source 1. .375
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0055"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 84.2%; Score 16; DB 7; Length 375;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 ATTATTGATTCTGTAG 17
              ||||||||||||||||
Db        312 ATTATTGATTCTGTAG 297
```

RESULT 24
BG360551
LOCUS BG360551 410 bp mRNA linear EST 01-DEC-2001
DEFINITION RTI00001 Rainbow Trout Intestine Lambda Zap II Vector library Oncorhynchus mykiss cDNA 5' similar to Unknown, mRNA sequence.
ACCESSION BG360551
VERSION BG360551.1 GI:17223416
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 410)
AUTHORS Kim,S. and Killefer,J.
TITLE Analysis of Rainbow Trout (Oncorhynchus mykiss) Intestine Expressed Sequence Tags
JOURNAL Unpublished (2001)
COMMENT Contact: Killefer J
Division of Animal and Veterinary Sciences
West Virginia University
PO Box 6108, Morgantown, WV 26506-6108, USA
Tel: 304 293 2631
Fax: 304 293 3740
Email: jkillef@wvu.edu
Insert Length: 410 Std Error: 0.00
Seq primer: T3 primer.
FEATURES Location/Qualifiers
source 1. .410
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/cell_type="Smooth muscle"
/clone_lib="Rainbow Trout Intestine Lambda Zap II Vector library"
/note="Organ: Intestine; Vector: pBluescript SK(+/-); Site_1: EcorR I; Site_2: Xho I
TAG_SEQ=Not found"
ORIGIN

Query Match 84.2%; Score 16; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        365 TATTGATTCTGTAGGT 380
```

RESULT 25
AI760729
LOCUS AI760729 457 bp mRNA linear EST 20-DEC-1999
DEFINITION wi08f05.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2389665 3', mRNA sequence.
ACCESSION AI760729
VERSION AI760729.1 GI:5176396
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 457)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever, M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 2291 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 443.
FEATURES Location/Qualifiers
source 1. .457
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2389665"
/tissue_type="B-cell, chronic lymphotic leukemia"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_CLL1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATTGCTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

Query Match 84.2%; Score 16; DB 1; Length 457;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 TTATTGATTCTGTAGG 18
              ||||||||||||||||
Db          4 TTATTGATTCTGTAGG 19
```

RESULT 26
CA061810/c
LOCUS CA061810 506 bp mRNA linear EST 04-MAR-2003
DEFINITION ssalrgp001052 mixed_tissue Salmo salar cDNA, mRNA sequence.
ACCESSION CA061810
VERSION CA061810.1 GI:24392053
KEYWORDS EST.
SOURCE Salmo salar (Atlantic salmon)
ORGANISM Salmo salar
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE 1 (bases 1 to 506)
AUTHORS GRASP Consortium, Davidson,W.S., Koop,B.F. and http://web.uvic.ca/cbr/grasp.
TITLE A survey of Salmo salar transcripts from high complexity cDNA libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada

```
            Tel: 250 472 4067
            Fax: 250 472 4075
            Email: bkoop@uvic.ca
            Centre for Biomedical Research, University of Victoria
            cDNA preparation and sequencing:
            Roberto Alberto, Marianne Beetz-Sargent, Maura Busby,
            Peter Hunt, Linda McKinnel, BF Koop.
            bioinformatics:
            Gordon D Brown
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .506
                      /organism="Salmo salar"
                      /mol_type="mRNA"
                      /strain="McConnell"
                      /db_xref="taxon:8030"
                      /clone_lib="mixed tissue"
                      /note="Vector: pCMVsport6; Library Creator:  Research
                      Genetics ; Atlantic salmon tissue contributors: Carlo
                      Biagi, Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.),
                      Simon Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery
                      (Crofton, B.C.), Rachel Roper (University of Victoria)"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 6;  Length 506;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        481 TATTGATTCTGTAGGT 466


RESULT 27
CO228348/c
LOCUS       CO228348                562 bp    mRNA    linear   EST 22-JUN-2004
DEFINITION  WS0014.B21_I04 WS-ES-A-1 Picea glauca cDNA clone WS0014_I04 3',
            mRNA sequence.
ACCESSION   CO228348
VERSION     CO228348.1  GI:49050663
KEYWORDS    EST.
SOURCE      Picea glauca (white spruce)
  ORGANISM  Picea glauca
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   Ralph,S., Kolosova,N., Cooper,D., Butterfield,Y., Kirkpatrick,R.,
            Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G., Babakaiff,R.,
            Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M.,
            Moran,J., Olson,T., Wong,D., Friedmann,M.F., Ritland,C.E.,
            Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C.,
            Ritland,K. and Bohlmann,J.
  TITLE     The spruce transcriptome: Analysis of expressed sequence tags from
            multiple cDNA libraries
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Joerg Bohlmann
            Genome BC forest genomics program
            University of British Columbia
            UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,
            Vancouver, British Columbia, Canada, V6T 1Z3
            Tel: 1-604-822-0282
            Fax: 1-604-822-6097
            Email: bohlmann@interchange.ubc.ca
            Plate: WS0014  row: I  column: 04
            High quality sequence stop: 562
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .562
                      /organism="Picea glauca"
                      /mol_type="mRNA"
                      /cultivar="PG-29"
                      /db_xref="taxon:3330"
                      /clone="WS0014_I04"
                      /sex="Hermaphrodite"
                      /tissue_type="Young shoots"
                      /lab_host="E. coli DH10B cells"
                      /clone_lib="WS-ES-A-1"
                      /note="Organ: Foliage from 25 year old trees harvested at
                      Kalamalka Research Station in Vernon, British Columbia on
                      June 15th, 2001; Vector: pBluescript II SK (+) XR; Site_1:
                      EcoRI (5' end of cDNA); Site_2: XhoI (3' end of cDNA);
                      cDNA was prepared from 5 micrograms of mRNA and
                      directionally ligated into the pBluescript II SK (+) XR
                      vector using the pBluescript II XR cDNA Library
                      Construction Kit according to manufacturer's instructions
                      with modifications (Stratagene). Plasmid DNA was then
                      transformed by electroporation into DH10B cells
                      (Invitrogen) for propagation."
ORIGIN

  Query Match             84.2%;  Score 16;  DB 7;  Length 562;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTA 16
              ||||||||||||||||
Db        126 AATTATTGATTCTGTA 111


RESULT 28
CK881762
LOCUS       CK881762                566 bp    mRNA    linear   EST 09-MAR-2004
DEFINITION  SGP142112 Atlantic salmon Head kidney cDNA library Salmo salar cDNA
            clone FN4-1269, mRNA sequence.
ACCESSION   CK881762
VERSION     CK881762.1  GI:45311393
KEYWORDS    EST.
SOURCE      Salmo salar (Atlantic salmon)
  ORGANISM  Salmo salar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Hoyheim,B.
  TITLE     Expressed Sequence Tags from an Atlantic salmon Head kidney cDNA
            library
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Bjorn Hoyheim
            Department of Basic Sciences and Aquatic Medicine
            Norwegian School of Veterinary Science
            PO Box 8146 DEP, NO-0033 Oslo, Norway
            Tel: 47 22 96 47 03
            Fax: 47 22 96 47 58
            Email: Bjorn.Hoyheim@veths.no.
FEATURES              Location/Qualifiers
     source           1. .566
                      /organism="Salmo salar"
                      /mol_type="mRNA"
                      /db_xref="taxon:8030"
                      /clone="FN4-1269"
                      /tissue_type="Head kidney"
                      /dev_stage="Pre-smolt"
                      /lab_host="XL 10-Gold"
                      /clone_lib="Atlantic salmon Head kidney cDNA library"
                      /note="Vector: pBluescript II SK (+) XR; Site_1: EcoRI;
                      Site_2: XhoI"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 7;  Length 566;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        146 TATTGATTCTGTAGGT 161
```

```
RESULT 29
AW912973/c
LOCUS       AW912973                581 bp    mRNA    linear   EST 25-MAY-2000
DEFINITION  uf47a10.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1514490 5' similar to gb:M74525 UBIQUITIN-CONJUGATING ENZYME
            E2-17 KD (HUMAN);, mRNA sequence.
ACCESSION   AW912973
VERSION     AW912973.1  GI:8078597
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: uf47a10.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:941342
            Seq primer: -40RP from Gibco
            High quality sequence stop: 332.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1514490"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             84.2%;  Score 16;  DB 2;  Length 581;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGG 18
              ||||||||||||||||
Db        293 TTATTGATTCTGTAGG 278


RESULT 30
BI946345
LOCUS       BI946345                581 bp    mRNA    linear   EST 19-OCT-2001
DEFINITION  01088 leafy spurge Lambda HybriZAP 2.1 two-hybrid vector cDNA
            Library Euphorbia esula cDNA clone 5AN 5' similar to tuber-specific
            and sucrose-responsive element binding factor, mRNA sequence.
ACCESSION   BI946345
VERSION     BI946345.1  GI:16284932
KEYWORDS    EST.
SOURCE      Euphorbia esula (leafy spurge)
  ORGANISM  Euphorbia esula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
            Euphorbieae; Euphorbia.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Anderson,J.V. and Horvath,D.P.
  TITLE     Identification of mRNAs expressed in underground adventitious buds
            of Euphorbia esula (leafy spurge)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Anderson JV
            Plants Science Research
            USDA/ARS, Biosciences Research Lab
            1605 Albrecht Blvd., PO Box 5674, Fargo, ND 58105, USA
            Tel: 701 239 1263
            Fax: 701 239 1252
            Email: andersjv@fargo.ars.usda.gov
            Seq primer: pAD5.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Euphorbia esula"
                     /mol_type="mRNA"
                     /db_xref="taxon:3993"
                     /clone="5AN"
                     /tissue_type="underground adventitious buds"
                     /dev_stage="3-day induced (decapitated)"
                     /clone_lib="leafy spurge Lambda HybriZAP 2.1 two-hybrid
                     vector cDNA Library"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 4;  Length 581;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTA 16
              ||||||||||||||||
Db        133 AATTATTGATTCTGTA 148


RESULT 31
CR315576
LOCUS       CR315576                583 bp    DNA     linear   GSS 01-MAR-2004
DEFINITION  Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
            truncatula, genomic survey sequence.
ACCESSION   CR315576
VERSION     CR315576.1  GI:44861720
KEYWORDS    GSS.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /cultivar="Jemalong A17"
                     /db_xref="taxon:3880"
                     /clone_lib="MTE1"
                     /note="Vector: pIndigoBAC ; Site_1: EcoRI ; Site_2: EcoRI
                     ; Debelle F. and Chalhoub B.~Genoscope sequence ID :
                     mte1-4G23RM1"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 9;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        306 TATTGATTCTGTAGGT 321
```

```
RESULT 32
CA043357/c
LOCUS       CA043357                 591 bp    mRNA    linear   EST 04-MAR-2003
DEFINITION  ssalplnb506296 gut Salmo salar cDNA, mRNA sequence.
ACCESSION   CA043357
VERSION     CA043357.1  GI:24344277
KEYWORDS    EST.
SOURCE      Salmo salar (Atlantic salmon)
  ORGANISM  Salmo salar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   GRASP Consortium, Davidson,W.S., Koop,B.F. and
            http://web.uvic.ca/cbr/grasp.
  TITLE     A survey of Salmo salar transcripts from high complexity cDNA
            libraries
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Koop BF
            Centre for Biomedical Research
            University of Victoria
            PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
            Tel: 250 472 4067
            Fax: 250 472 4075
            Email: bkoop@uvic.ca
            Genome Sciences Centre, BC Cancer Agency
            cDNA preparation, sequencing and bioinformatics:
            Y Butterfield, R Kirkpatrick, J Asano, N Girn, R Guin, D Lee,
            S Lee, T Olson, P Pandoh, A Prahbu, D Smailus, L Spence, J Stott,
            S Taylor, G Yang, J Schein, S Jones and M Marra.
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="Salmo salar"
                     /mol_type="mRNA"
                     /strain="McConnell"
                     /db_xref="taxon:8030"
                     /clone_lib="gut"
                     /note="Vector: pBlueScriptIISK+; Library Creator:  Matthew
                     L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,
                     Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.), Simon
                     Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery (Crofton,
                     B.C.), Rachel Roper (University of Victoria)"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 6;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        432 TATTGATTCTGTAGGT 417


RESULT 33
CD334984
LOCUS       CD334984                 611 bp    mRNA    linear   EST 17-SEP-2003
DEFINITION  StrPu536.001806 Sea urchin embryo 40hr gastrula stage cDNA library
            MPMGp536 Strongylocentrotus purpuratus cDNA clone
            CALTp536A1138;MPI_536_38A11 5', mRNA sequence.
ACCESSION   CD334984
VERSION     CD334984.1  GI:34801507
KEYWORDS    EST.
SOURCE      Strongylocentrotus purpuratus
  ORGANISM  Strongylocentrotus purpuratus
            Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
            Echinoidea; Euechinoidea; Echinacea; Echinoida;
            Strongylocentrotidae; Strongylocentrotus.
REFERENCE   1  (bases 1 to 611)
  AUTHORS   Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,
            Reinhardt,R., Herwig,R., Panopoulou,G. and Lehrach,H.
  TITLE     Generation, annotation, evolutionary analysis, and database
            integration of 20,000 unique sea urchin EST clusters
  JOURNAL   Genome Res. 13 (12), 2736-2746 (2003)
COMMENT     Contact: Poustka AJ
            laboraty 145, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1235
            Fax: +49 30 8413 1128
            Email: poustka@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONF) to reduce sequencing redundancy. According to the ONF
            procedure, clones that display the same hybridisation matrix with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per ONF cluster is selected for sequencing. The size of each
            cluster is an indicator of the frequency of a transcript in the
            analysed library. The cluster size as well as the coordinates of
            the other clones assigned to the same ONF cluster as the clone from
            which the above EST is generated is available at the sea urchin
            project web site at: http://www.molgen.mpg.de/ag_seaurchin/ . cDNA
            clones and filters are distributed via the Resource Center/Primary
            Database of the German Human Genome Project (http://www.rzpd.de)
            PCR PRimers
            FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP) 5'-seq
            BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP) 3'-seq
            Seq primer: 5'-CCGGTCCGGAATTCCCGGGT-3' pSport3/86
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1. .611
                     /organism="Strongylocentrotus purpuratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:7668"
                     /clone="CALTp536A1138;MPI_536_38A11"
                     /tissue_type="whole embryo"
                     /dev_stage="embryonic 40hr"
                     /lab_host="E.coli, XL1 blue"
                     /clone_lib="Sea urchin embryo 40hr gastrula stage cDNA
                     library MPMGp536"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI; Random
                     primed and directionally cloned in pSport1 vector using a
                     NotI (5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15-3' and a
                     SalI 5'- TCGACCCACGCGTCCG-3'adapters (Gibco BRL)"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 6;  Length 611;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        297 TATTGATTCTGTAGGT 312


RESULT 34
CA353561
LOCUS       CA353561                 622 bp    mRNA    linear   EST 05-NOV-2002
DEFINITION  625083 NCCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT69P08_D_H04 5',
            mRNA sequence.
ACCESSION   CA353561
VERSION     CA353561.1  GI:24598732
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   Rexroad,C.E. 3rd, Lee,Y., Keele,J.W., Karamycheva,S., Brown,G.,
            Koop,B., Gahr,S.A., Palti,Y. and Quackenbush,J.
  TITLE     Sequence analysis of a rainbow trout cDNA library and creation of a
            gene index
  JOURNAL   Cytogenet. Genome Res. 102 (1-4), 347-354 (2003)
COMMENT     Contact: Rexroad CE
            USDA, ARS, National Center for Cool and Cold Water Aquaculture
            11876 Leetown Road, Kearneysville, WV 25430, USA
```

```
            Tel: 304 724 8340 x2129
            Fax: 304 725 0351
            Email: crexroad@ncccwa.ars.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified by
            cross_match v0.990329.
            Seq primer: AGCGGATAACAATTTCACACAGGA.
FEATURES              Location/Qualifiers
     source           1. .622
                      /organism="Oncorhynchus mykiss"
                      /mol_type="mRNA"
                      /db_xref="taxon:8022"
                      /clone="1RT69P08_D_H04"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /clone_lib="NCCCWA 1RT"
                      /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                      Library made from pooled tissue from brain, gill, liver,
                      spleen, muscle, and kidney."
ORIGIN

  Query Match             84.2%;  Score 16;  DB 6;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        400 TATTGATTCTGTAGGT 415


RESULT 35
CO217921/c
LOCUS       CO217921                634 bp    mRNA    linear   EST 22-JUN-2004
DEFINITION  WS0101.B21_K21 SS-R-N-A-11 Picea sitchensis cDNA clone WS0101_K21
            3', mRNA sequence.
ACCESSION   CO217921
VERSION     CO217921.1  GI:49040235
KEYWORDS    EST.
SOURCE      Picea sitchensis (Sitka spruce)
  ORGANISM  Picea sitchensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Ralph,S., Kolosova,N., Cooper,D., Butterfield,Y., Kirkpatrick,R.,
            Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G., Babakaiff,R.,
            Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M.,
            Moran,J., Olson,T., Wong,D., Friedmann,M.F., Ritland,C.E.,
            Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C.,
            Ritland,K. and Bohlmann,J.
  TITLE     The spruce transcriptome: Analysis of expressed sequence tags from
            multiple cDNA libraries
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Joerg Bohlmann
            Genome BC forest genomics program
            University of British Columbia
            UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,
            Vancouver, British Columbia, Canada, V6T 1Z3
            Tel: 1-604-822-0282
            Fax: 1-604-822-6097
            Email: bohlmann@interchange.ubc.ca
            Plate: WS0101  row: K  column: 21
            High quality sequence stop: 634
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .634
                      /organism="Picea sitchensis"
                      /mol_type="mRNA"
                      /cultivar="Gb2-229"
                      /db_xref="taxon:3332"
                      /clone="WS0101_K21"
                      /sex="Hermaphrodite"
                      /tissue_type="Young root growth (terminal 1-3 cm) and old
                      root growth (distal to terminal 1-3 cm) tissues"
                      /dev_stage="three year old clonal trees grown under
                      greenhouse conditions in standard potting soil mixture."
                      /lab_host="E. coli DH10B cells"
                      /clone_lib="SS-R-N-A-11"
                      /note="Organ: Roots; Vector: pBluescript II SK (+) XR;
                      Site_1: EcoRI (5' end of cDNA); Site_2: XhoI (3' end of
                      cDNA); mRNA was isolated from each tissue source
                      independently and equal quantities of mRNA from each
                      tissue were then pooled. cDNA was prepared from 5
                      micrograms of mRNA and directionally ligated into the
                      pBluescript II SK (+) XR vector using the pBluescript II
                      XR cDNA Library Construction Kit according to
                      manufacturer's instructions with modifications
                      (Stratagene). Plasmid DNA was then transformed by
                      electroporation into DH10B cells (Invitrogen) for
                      propagation. Normalization was applied according to
                      published methods [Bonaldo M.F. et al. (1996) Genome
                      Research 6(9):791] in order to reduce the abundance of
                      highly expressed transcripts."
ORIGIN

  Query Match             84.2%;  Score 16;  DB 7;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTA 16
              ||||||||||||||||
Db        139 AATTATTGATTCTGTA 124


RESULT 36
CB502289
LOCUS       CB502289                642 bp    mRNA    linear   EST 16-MAY-2003
DEFINITION  ssalplnb506296_rev gut Salmo salar cDNA, mRNA sequence.
ACCESSION   CB502289
VERSION     CB502289.1  GI:29313515
KEYWORDS    EST.
SOURCE      Salmo salar (Atlantic salmon)
  ORGANISM  Salmo salar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   GRASP Consortium, Davidson,W.S., Koop,B.F. and
            http://web.uvic.ca/cbr/grasp.
  TITLE     A survey of Salmo salar transcripts from high complexity cDNA
            libraries
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Koop BF
            Centre for Biomedical Research
            University of Victoria
            PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
            Tel: 250 472 4067
            Fax: 250 472 4075
            Email: bkoop@uvic.ca
            Genome Sciences Centre, BC Cancer Agency cDNA preparation,
            sequencing and bioinformatics: Y Butterfield, R Kirkpatrick, J
            Asano, N Girn, R Guin, D Lee, S Lee, T Olson, P Pandoh, A Prahbu, D
            Smailus, L Spence, J Stott, S Taylor, G Yang, J Schein, S Jones and
            M Marra.
FEATURES              Location/Qualifiers
     source           1. .642
                      /organism="Salmo salar"
                      /mol_type="mRNA"
                      /strain="McConnell"
                      /db_xref="taxon:8030"
                      /clone_lib="gut"
                      /note="Vector: pBlueScriptIISK+; Library Creator:  Matthew
                      L Rise ; Atlantic salmon tissue contributors: Carlo Biagi,
                      Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.), Simon
                      Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery (Crofton,
                      B.C.), Rachel Roper (University of Victoria)"
ORIGIN
```

```
  Query Match             84.2%;  Score 16;  DB 6;  Length 642;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        627 TATTGATTCTGTAGGT 642


RESULT 37
AG151770
LOCUS       AG151770                  647 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-014P17.TJ, genomic survey
            sequence.
ACCESSION   AG151770
VERSION     AG151770.1  GI:16681448
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 647)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: TJ
             LIBRARY
                 Vector     : pBACe3.6
                 R.Site 1   : EcoRI
                 R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-014P17.TJ"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 9;  Length 647;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTA 16
              ||||||||||||||||
Db        496 AATTATTGATTCTGTA 511


RESULT 38
CA377750
LOCUS       CA377750                  648 bp    mRNA    linear   EST 06-NOV-2002
DEFINITION  656393 NCCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT40D13_B_B07 5',
            mRNA sequence.
ACCESSION   CA377750
VERSION     CA377750.1  GI:24697047
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Rexroad,C.E. 3rd, Lee,Y., Keele,J.W., Karamycheva,S., Brown,G.,
            Koop,B., Gahr,S.A., Palti,Y. and Quackenbush,J.
  TITLE     Sequence analysis of a rainbow trout cDNA library and creation of a
            gene index
  JOURNAL   Cytogenet. Genome Res. 102 (1-4), 347-354 (2003)
COMMENT     Contact: Rexroad CE
            USDA, ARS, National Center for Cool and Cold Water Aquaculture
            11876 Leetown Road, Kearneysville, WV 25430, USA
            Tel: 304 724 8340 x2129
            Fax: 304 725 0351
            Email: crexroad@ncccwa.ars.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified by
            cross_match v0.990329.
            Seq primer: AGCGGATAACAATTTCACACAGGA.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /db_xref="taxon:8022"
                     /clone="1RT40D13_B_B07"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="NCCCWA 1RT"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from brain, gill, liver,
                     spleen, muscle, and kidney."
ORIGIN

  Query Match             84.2%;  Score 16;  DB 6;  Length 648;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        206 TATTGATTCTGTAGGT 221


RESULT 39
BX077071
LOCUS       BX077071                  651 bp    mRNA    linear   EST 07-MAY-2004
DEFINITION  BX077071 AGENAE Rainbow trout normalized multi-tissues library
            (tcad) Oncorhynchus mykiss cDNA clone tcad0005a.1.16 3prim, mRNA
            sequence.
ACCESSION   BX077071
VERSION     BX077071.2  GI:42604961
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 651)
  AUTHORS   Govoroun,M., Guiguen,Y. and Le Gac,F.
  TITLE     Construction and primary characterization of normalized cDNA
            libraries in rainbow trout, Oncorhynchus mykiss
  JOURNAL   Unpublished (2003)
COMMENT     On Jan 14, 2003 this sequence version replaced gi:27740490.
            Contact: Guiguen Y
            INRA - SCRIBE
            Campus de beaulieu, RENNES cedex, 35042, France
            Tel: 02.23.48.50.09
            Fax: 02.23.48.50.20
            Email: Yann.Guiguen@beaulieu.rennes.inra.fr
            Sequence cleaned of vector, adaptator and repetitions. Contact us
            at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
            sequence.
```

```
           Plate: 0005  row: 1  column: 16
           Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .651
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /db_xref="taxon:8022"
                     /clone="tcad0005a.1.16"
                     /tissue_type="adipose tissue, blood, brain,
                     differentiating gonads, interrenal, intestine, kidney,
                     liver, muscle, ovary, pituitary, testis"
                     /lab_host="DH10B"
                     /clone_lib="AGENAE Rainbow trout normalized multi-tissues
                     library (tcad)"
                     /note="Vector: pT7T3D-pac; Clone distribution : AGENAE
                     Resource centre. Francois PIUMI,
                     Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
                     Etude du genome (LREG), Domaine de Vilvert, 78352,
                     Jouy-en-Josas cedex, France, +33 (0) 1.34.65.28.02, +33
                     (0) 1.34.65.22.73"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 5;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        147 TATTGATTCTGTAGGT 162


RESULT 40
CE160496/c
LOCUS       CE160496                651 bp    DNA     linear   GSS 25-SEP-2003
DEFINITION  tigr-gss-dog-17000371420074 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE160496
VERSION     CE160496.1  GI:35289890
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 651)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .651
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 9;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        129 TATTGATTCTGTAGGT 114


RESULT 41
CB497718/c
LOCUS       CB497718                655 bp    mRNA    linear   EST 27-MAR-2003
DEFINITION  omykrbna501060 Oncorhynchus mykiss Chilliwack River steelhead whole
            Oncorhynchus mykiss cDNA, mRNA sequence.
ACCESSION   CB497718
VERSION     CB497718.1  GI:29308944
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   GRASP Consortium, Davidson,W.S., Koop,B.F. and
            http://web.uvic.ca/cbr/grasp.
  TITLE     A survey of Salmo salar transcripts from high complexity cDNA
            libraries
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Koop BF
            Centre for Biomedical Research
            University of Victoria
            PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
            Tel: 250 472 4067
            Fax: 250 472 4075
            Email: bkoop@uvic.ca
            Genome Sciences Centre, BC Cancer Agency cDNA preparation,
            sequencing and bioinformatics: Y Butterfield, R Kirkpatrick, J
            Asano, N Girn, R Guin, D Lee, S Lee, T Olson, P Pandoh, A Prahbu, D
            Smailus, L Spence, J Stott, S Taylor, G Yang, J Schein, S Jones and
            M Marra.
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /strain="Chilliwack River steelhead"
                     /db_xref="taxon:8022"
                     /clone_lib="Oncorhynchus mykiss Chilliwack River steelhead
                     whole"
                     /note="Vector: pBlueScriptIISK+; Library Creator:  Matthew
                     L Rise, BF Koop ; Rainbow trout tissue contributors:
                     Robert Devlin (DFO, Vancouver, B.C.)"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 6;  Length 655;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        495 TATTGATTCTGTAGGT 480


RESULT 42
CE545000/c
LOCUS       CE545000                669 bp    DNA     linear   GSS 28-SEP-2003
DEFINITION  tigr-gss-dog-17000366094583 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE545000
VERSION     CE545000.1  GI:36861781
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 669)
```

```
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             84.2%;  Score 16;  DB 9;  Length 669;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        554 TATTGATTCTGTAGGT 539


RESULT 43
CO204386/c
LOCUS       CO204386                683 bp    mRNA    linear   EST 21-JUN-2004
DEFINITION  WS0064.B21_F14 IS-B-A-7 Picea engelmannii x Picea sitchensis cDNA
            clone WS0064_F14 3', mRNA sequence.
ACCESSION   CO204386
VERSION     CO204386.1  GI:49015561
KEYWORDS    EST.
SOURCE      Picea engelmannii x Picea sitchensis
  ORGANISM  Picea engelmannii x Picea sitchensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE   1  (bases 1 to 683)
  AUTHORS   Ralph,S., Kolosova,N., Cooper,D., Butterfield,Y., Kirkpatrick,R.,
            Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G., Babakaiff,R.,
            Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M.,
            Moran,J., Olson,T., Wong,D., Friedmann,M.F., Ritland,C.E.,
            Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C.,
            Ritland,K. and Bohlmann,J.
  TITLE     The spruce transcriptome: Analysis of expressed sequence tags from
            multiple cDNA libraries
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Joerg Bohlmann
            Genome BC forest genomics program
            University of British Columbia
            UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,
            Vancouver, British Columbia, Canada, V6T 1Z3
            Tel: 1-604-822-0282
            Fax: 1-604-822-6097
            Email: bohlmann@interchange.ubc.ca
            Plate: WS0064  row: F  column: 14
            High quality sequence stop: 683
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Picea engelmannii x Picea sitchensis"
                     /mol_type="mRNA"
                     /cultivar="Fal-1028"
                     /db_xref="taxon:273280"
                     /clone="WS0064_F14"
                     /sex="Hermaphrodite"
                     /lab_host="E. coli DH10B cells"
                     /clone_lib="IS-B-A-7"
                     /note="Organ: Bark (with phloem and cambium attached) from
                     one year old clonal trees grown under greenhouse
                     conditions in standard potting soil mixture; Vector:
                     pBluescript II SK (+) XR; Site_1: EcoRI (5' end of cDNA);
                     Site_2: XhoI (3' end of cDNA); Bark was wounded using
                     razor blades along the entire length of the tree at 5 mm
                     intervals on opposite sides of the trunk. The same trees
                     were also sprayed with a 0.01% (v/v) methyl jasmonate
                     solution resuspended in 0.1% (v/v) tween 20 (~50mLs per
                     tree). Bark tissue with phloem attached was harvested 24
                     hours, 2 days, 4 days and 8 days after initiating the
                     treatment. mRNA was isolated from each tissue source
                     independently and equal quantities of mRNA from each
                     tissue were then pooled. cDNA was prepared from 5
                     micrograms of mRNA and directionally ligated into the
                     pBluescript II SK (+) XR vector using the pBluescript II
                     XR cDNA Library Construction Kit according to
                     manufacturer's instructions with modifications
                     (Stratagene). Plasmid DNA was then transformed by
                     electroporation into DH10B cells (Invitrogen) for
                     propagation."
ORIGIN

  Query Match             84.2%;  Score 16;  DB 7;  Length 683;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTA 16
              ||||||||||||||||
Db        303 AATTATTGATTCTGTA 288


RESULT 44
BH973406/c
LOCUS       BH973406                686 bp    DNA     linear   GSS 02-OCT-2002
DEFINITION  odj04e11.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
            sequence.
ACCESSION   BH973406
VERSION     BH973406.1  GI:23456409
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
            Nash,W., Rabinowicz,P.D. and Wilson,R.K.
  TITLE     Whole genome shotgun reads from Brassica oleracea
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Plate: odj04  row: e  column: 11
            Seq primer: -28RPpOT reverse
            Class: shotgun
            High quality sequence start: 24
            High quality sequence stop: 457.
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3712"
                     /clone_lib="B.oleracea002"
                     /note="Vector: pOTw13; Whole genome shotgun library from
                     flowering buds. DNA was purified  from a crude nuclear
                     prep using Brassica oleracea TO1000DH3 buds  provided by
```

Thosmas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 84.2%; Score 16; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db         95 TATTGATTCTGTAGGT 80
```

RESULT 45
CO240857/c
LOCUS CO240857 687 bp mRNA linear EST 22-JUN-2004
DEFINITION WS0078.B21_H05 WS-PS-N-A-8 Picea glauca cDNA clone WS0078_H05 3', mRNA sequence.
ACCESSION CO240857
VERSION CO240857.1 GI:49063174
KEYWORDS EST.
SOURCE Picea glauca (white spruce)
ORGANISM Picea glauca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE 1 (bases 1 to 687)
AUTHORS Ralph,S., Kolosova,N., Cooper,D., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G., Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A., Mayo,M., Moran,J., Olson,T., Wong,D., Friedmann,M.F., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J.
TITLE The spruce transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
JOURNAL Unpublished (2004)
COMMENT Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-6097
Email: bohlmann@interchange.ubc.ca
Plate: WS0078 row: H column: 05
High quality sequence stop: 687
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .687
/organism="Picea glauca"
/mol_type="mRNA"
/cultivar="PG-29"
/db_xref="taxon:3330"
/clone="WS0078_H05"
/sex="Hermaphrodite"
/tissue_type="Flushing buds harvested May 16th, young shoots harvested June 15th, and mature shoots harvested June 15th"
/lab_host="E. coli DH10B cells"
/clone_lib="WS-PS-N-A-8"
/note="Organ: Foliage from 25 year old trees harvested at Kalamalka Research Station in Vernon, British Columbia in 2001; Vector: pBluescript II SK (+) XR; Site_1: EcoRI (5' end of cDNA); Site_2: XhoI (3' end of cDNA); mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Query Match 84.2%; Score 16; DB 7; Length 687;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTA 16
              ||||||||||||||||
Db        145 AATTATTGATTCTGTA 130
```

Search completed: April 26, 2005, 16:54:18
Job time : 3141 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 26, 2005, 08:28:02 ; Search time 436 Seconds
(without alignments)
257.970 Million cell updates/sec

Title: US-10-619-906-5
Perfect score: 19
Sequence: 1 aattattgattctgtaggt 19

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	19	100.0	19	12	ADI53692	Adi53692 Human MMP
c	2	19	100.0	458	6	ABK53984	Abk53984 Human hea
c	3	19	100.0	589	6	ABQ59213	Abq59213 Human col
c	4	19	100.0	1413	8	ABX76392	Abx76392 Lung canc
c	5	19	100.0	1413	10	ADF18051	Adf18051 Human mat
c	6	19	100.0	1413	11	ADN39005	Adn39005 Cancer/an
c	7	19	100.0	1413	11	ADN39708	Adn39708 Cancer/an
c	8	19	100.0	1413	11	ADN39709	Adn39709 Cancer/an
c	9	19	100.0	1653	6	ABV78217	Abv78217 Human MMP
c	10	19	100.0	1653	6	ABZ35793	Abz35793 Human MMP
c	11	19	100.0	1653	6	ABX10036	Abx10036 Human MMP
c	12	19	100.0	1653	6	ABL91758	Abl91758 Human pol
c	13	19	100.0	1778	3	AAA35052	Aaa35052 Human ade
c	14	19	100.0	1778	3	AAF21174	Aaf21174 Human low
c	15	19	100.0	1778	4	AAH28229	Aah28229 Nucleotid
c	16	19	100.0	1778	6	ABL62664	Abl62664 Colon ade
c	17	19	100.0	1778	6	ABL65817	Abl65817 Lung canc
c	18	19	100.0	1778	6	ABL64743	Abl64743 Lung canc
c	19	19	100.0	1778	6	ABL62102	Abl62102 Colon ade
c	20	19	100.0	1778	6	ABL66489	Abl66489 Lung canc
c	21	19	100.0	1778	6	ABN95681	Abn95681 Gene #217
c	22	19	100.0	1778	6	ABK54023	Abk54023 Human hea
c	23	19	100.0	1778	8	ACC51017	Acc51017 Human bla
c	24	19	100.0	1778	8	ABX76137	Abx76137 Lung canc
c	25	19	100.0	1778	8	ACF12895	Acf12895 Human cer
c	26	19	100.0	1778	10	ADD18687	Add18687 Human dis
c	27	19	100.0	1778	10	ADF18049	Adf18049 cDNA enco
c	28	19	100.0	1778	10	ADF74444	Adf74444 Full leng
c	29	19	100.0	1778	10	ACC46771	Acc46771 Human COP
c	30	19	100.0	1778	10	ABZ96868	Abz96868 Human nuc
c	31	19	100.0	1778	11	ABD20717	Abd20717 Human pul
c	32	19	100.0	1778	12	ADI53688	Adi53688 Human mat
c	33	19	100.0	1778	12	ADN05808	Adn05808 Antipsori
c	34	19	100.0	1778	12	ADQ29584	Adq29584 Human col
c	35	19	100.0	1778	13	ADR24908	Adr24908 Breast ca
c	36	19	100.0	1781	8	ABZ20476	Abz20476 Matrix me
c	37	19	100.0	1818	4	AAF81624	Aaf81624 Human mac
c	38	19	100.0	1818	4	AAC64994	Aac64994 Human mac
c	39	19	100.0	1873	12	ADQ23366	Adq23366 Human sof
c	40	19	100.0	1874	3	AAC77981	Aac77981 Human can
c	41	19	100.0	1988	10	ADB47336	Adb47336 Human cDN
c	42	19	100.0	2870	10	ADD29857	Add29857 Human tum
c	43	19	100.0	9137	3	AAA35055	Aaa35055 Human ade
c	44	19	100.0	9137	3	AAF21177	Aaf21177 Human low
c	45	19	100.0	9137	10	ABZ96871	Abz96871 Human nuc

ALIGNMENTS

RESULT 1
ADI53692
ID ADI53692 standard; DNA; 19 BP.
XX
AC ADI53692;
XX
DT 22-APR-2004 (first entry)
XX
DE Human MMP-12 antisense oligonucleotide, SEQ ID 5.
XX
KW Antiinflammatory; Antipsoriatic; Antiasthmatic; Antiarthritic;
KW Respiratory; antisense oligonucleotide; matrix metalloproteinase 12;
KW MMP-12; inflammatory bowel disease; ulcerative colitis; Crohn's disease;
KW rheumatoid arthritis; psoriasis; emphysema; asthma; human; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2004009098-A1.
XX
PD 29-JAN-2004.
XX
PF 17-JUL-2003; 2003WO-SE001223.
XX
PR 18-JUL-2002; 2002SE-00002253.
PR 04-SEP-2002; 2002US-0407680P.
XX
PA (INDE-) INDEX PHARM AB.
XX
PI Dieckmann A, Loefberg R, Von Stein O, Von Stein P, Good L;
XX
DR WPI; 2004-123288/12.
XX
PT New compound having a sequence targeted to a nucleic acid encoding
PT metalloproteinase 12 (MMP-12), useful for preparing a composition for
PT treating or preventing MMP-12 dependent disorder in a human patient e.g.,
PT asthma or psoriasis.
XX
PS Claim 7; SEQ ID NO 5; 55pp; English.
XX
CC The present invention relates to antisense oligonucleotides (ADI53690-
CC ADI53701) for matrix metalloproteinase 12 (MMP-12; ADI53688 and
CC ADI53689), which specifically hybridise with the nucleic acid encoding

```
CC   MMP-12 and inhibiting the translation of MMP-12 protein. The antisense
CC   oligonucleotides are useful for preparing a composition for treating or
CC   preventing MMP-12 dependent disorder in a human patient e.g. inflammatory
CC   bowel disease, such as ulcerative colitis and Crohn's disease, rheumatoid
CC   arthritis, psoriasis, emphysema or asthma.
XX
SQ   Sequence 19 BP; 5 A; 1 C; 4 G; 9 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 12;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 0.35;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db          1 AATTATTGATTCTGTAGGT 19


RESULT 2
ABK53984/c
ID   ABK53984 standard; cDNA; 458 BP.
XX
AC   ABK53984;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Human head and neck tumour cDNA, SEQ ID No 179.
XX
KW   Human; head and neck cancer; tumour; cytostatic; immunogenic; vaccine;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200212329-A2.
XX
PD   14-FEB-2002.
XX
PF   01-AUG-2001; 2001WO-US024226.
XX
PR   03-AUG-2000; 2000US-0223281P.
PR   16-NOV-2000; 2000US-0249933P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Fan L;
XX
DR   WPI; 2002-257467/30.
XX
PT   Novel polynucleotide encoding head and neck tumor polypeptides, useful in
PT   pharmaceutical compositions, e.g. vaccines, for treating head and neck
PT   cancers.
XX
PS   Claim 1; Page 168-169; 200pp; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising
CC   sequences selected from 273 sequences fully defined in the specification.
CC   (I), including its encoded polypeptide (II), an antibody binding to (II),
CC   a fusion protein comprising (II) and a T-cell population stimulated by
CC   (I) or (II) are useful for stimulating an immune response in a patient
CC   and treating head and neck cancer in a patient. An oligonucleotide (III)
CC   that hybridises to (I) is useful for determining the presence of cancer
CC   in a patient, by obtaining a biological sample from the patient,
CC   contacting the sample with (III), detecting in the sample an amount of a
CC   polynucleotide that hybridises to the oligonucleotide, and comparing the
CC   amount of polynucleotide that hybridises to the oligonucleotides to a
CC   predetermined cut-off value. (I) and (II) are useful in pharmaceutical
CC   compositions, e.g. vaccines, and other compositions for the diagnosis and
CC   treatment of head and neck cancer. ABK53806-ABK54078 represent human head
CC   and neck cancer cDNA sequences of the invention
XX
SQ   Sequence 458 BP; 151 A; 99 C; 100 G; 108 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 6;  Length 458;
  Best Local Similarity   100.0%;  Pred. No. 0.34;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        348 AATTATTGATTCTGTAGGT 330


RESULT 3
ABQ59213/c
ID   ABQ59213 standard; cDNA; 589 BP.
XX
AC   ABQ59213;
XX
DT   02-AUG-2002  (first entry)
XX
DE   Human colon cancer related nucleotide sequence SEQ ID NO:2908.
XX
KW   Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW   genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200229086-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030732.
XX
PR   02-OCT-2000; 2000US-0237271P.
XX
PA   (FARB ) BAYER CORP.
XX
PI   Burgess C,  Astle JH,  Carroll E,  Catino TJ,  Dwivedi P,  Molino GA;
PI   Thiaglingam A,  Lewis ME;
XX
DR   WPI; 2002-426115/45.
XX
PT   New isolated nucleic acid that is differentially expressed in cancer
PT   tissues useful for determining the presence of colon cancer in a cell or
PT   tissue type, and in antisense therapy.
XX
PS   Claim 1; Fig 1; 796pp; English.
XX
CC   ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC   expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC   encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC   used in antisense therapy. An antibody immunoreactive with a polypeptide
CC   encoded by (I) is useful for detecting cancer in a patient sample, and
CC   for detecting the presence or absence of a polynucleotide encoded by a
CC   nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC   from (I) can be used for determining the presence of a nucleic acid which
CC   hybridises to (I), and for determining the phenotype of cells in a sample
CC   of cells from a patient. (I) is useful for determining the presence of
CC   colon cancer in a cell or tissue type, for determining the presence or
CC   state of other type of cancer, in antisense therapy, to generate
CC   macroarrays on a solid surface, to identify a chromosome on which the
CC   corresponding gene resides, and in tissue profiling, forensics, genetic
CC   analysis, mapping and diagnostic applications. (I) can be used to raise
CC   antibodies, and to screen for peptide analogues and antagonists
XX
SQ   Sequence 589 BP; 182 A; 120 C; 134 G; 144 T; 0 U; 9 Other;

  Query Match             100.0%;  Score 19;  DB 6;  Length 589;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        419 AATTATTGATTCTGTAGGT 401


RESULT 4
ABX76392/c
```

```
ID   ABX76392 standard; DNA; 1413 BP.
XX
AC   ABX76392;
XX
DT   02-APR-2003  (first entry)
XX
DE   Lung cancer-associated polynucleotide #256.
XX
KW   Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW   antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW   small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW   chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW   interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS   Unidentified.
XX
PN   WO200286443-A2.
XX
PD   31-OCT-2002.
XX
PF   18-APR-2002; 2002WO-US012476.
XX
PR   18-APR-2001; 2001US-0284770P.
PR   10-MAY-2001; 2001US-0290492P.
PR   09-NOV-2001; 2001US-0339245P.
PR   13-NOV-2001; 2001US-0350666P.
PR   29-NOV-2001; 2001US-0334370P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Aziz N,  Murray R;
XX
DR   WPI; 2003-093161/08.
DR   P-PSDB; ABU56663.
XX
PT   Detecting a lung cancer-associated transcript in a cell from a patient
PT   for treating lung cancer, by contacting a biological sample from the
PT   patient with a polynucleotide that exhibits increased or decreased
PT   expression in lung cancer.
XX
PS   Claim 22; Page 385; 453pp; English.
XX
CC   The invention relates to a method for detecting a lung cancer-associated
CC   transcript in a cell from a patient, comprising contacting a biological
CC   sample from the patient with a polynucleotide that selectively hybridises
CC   to a sequence that is at least 80 % identical to a gene that exhibits
CC   increased or decreased expression in lung cancer samples. Lung cancer-
CC   associated polynucleotides and polypeptides are used for identifying a
CC   compound that modulates a lung cancer-associated polypeptide, for
CC   inhibiting proliferation of a lung cancer-associated cell to treat lung
CC   cancer in a patient and for treating a mammal having lung cancer by
CC   administering a modulatory compound identified. The methods are useful
CC   for treating lung cancer, such as small cell lung cancer, non-small cell
CC   lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC   emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC   hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC   bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC   for diagnostic purposes and as targets for screening for therapeutic
CC   compounds that modulate lung cancer, such as antibodies. Sequences
CC   ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC   invention
XX
SQ   Sequence 1413 BP; 422 A; 306 C; 289 G; 396 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 8;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343
```

```
RESULT 5
ADF18051/c
ID   ADF18051 standard; cDNA; 1413 BP.
XX
AC   ADF18051;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human matrix metalloproteinase 12 (MMP-12) coding sequence.
XX
KW   anoretic; antidiabetic; immunomodulator; MMP-12 agonist;
KW   MMP-12 antagonist; MMP-12 inverse agonist; anti-MMP-12 antibody;
KW   antisense MMP-12; MMP-12 ribozyme; lipid homeostasis modulator;
KW   glucose homeostasis modulator; insulin homeostasis modulator;
KW   adipocyte growth modulator; metabolic disorder;
KW   matrix metalloproteinase 12; MMP-12; translation inhibition; obesity;
KW   overweight; diabetes; insulin resistance; cachexia; anorexia; gene; ss;
KW   human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1413
FT                   /*tag=  a
FT                   /product= "Human MMP-12"
FT                   /note= "Matrix metalloproteinase"
XX
PN   US2003157110-A1.
XX
PD   21-AUG-2003.
XX
PF   07-JAN-2003; 2003US-00337632.
XX
PR   07-JAN-2002; 2002US-0346354P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   An WF,  Chen H;
XX
DR   WPI; 2003-897814/82.
DR   P-PSDB; ADF18050.
XX
PT   Identifying a compound capable of treating a metabolic disorder comprises
PT   assaying for compounds that modulate matrix metalloprotease MMP-12
PT   expression or activity.
XX
PS   Disclosure; SEQ ID NO 3; 38pp; English.
XX
CC   The invention describes a method of identifying a compound capable of
CC   treating a metabolic disorder comprising assaying the ability of the
CC   compound to modulate a matrix metalloproteinase 12 (MMP-12) nucleic acid
CC   expression or MMP-12 polypeptide activity. The compound has anoretic,
CC   antidiabetic and immunomodulator properties. An antisense oligonucleotide
CC   or ribozyme targets the MMP-12 transcripts and inhibits translation.
CC   Alternatrively the oligonucleotide forms a triple helix with the promoter
CC   of the MMP-12 gene and inhibits transcription. The invention is used to
CC   identify a compound useful for treating a metabolic disorder,
CC   particularly obesity, overweight, diabetes, insulin resistance, cachexia
CC   or anorexia. This sequence encodes human matrix metalloproteinase 12 (MMP
CC   -12).
XX
SQ   Sequence 1413 BP; 421 A; 306 C; 289 G; 397 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343
```

RESULT 6
ADN39005/c
ID ADN39005 standard; cDNA; 1413 BP.
XX
AC ADN39005;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:323.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-PSDB; ADN39006.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO 323; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularistaion syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 1413 BP; 422 A; 306 C; 289 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 11; Length 1413;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 361 AATTATTGATTCTGTAGGT 343

RESULT 7
ADN39708/c
ID ADN39708 standard; cDNA; 1413 BP.
XX
AC ADN39708;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:C80.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-PSDB; ADN39925.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO C80; 1385pp; English.

XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularistaion syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 1413 BP; 422 A; 306 C; 289 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 11; Length 1413;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 361 AATTATTGATTCTGTAGGT 343

RESULT 8
ADN39709/c
ID ADN39709 standard; cDNA; 1413 BP.
XX
AC ADN39709;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:C81.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-PSDB; ADN39926.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO C81; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularistaion syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 1413 BP; 421 A; 308 C; 288 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 11; Length 1413;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 361 AATTATTGATTCTGTAGGT 343

RESULT 9
ABV78217/c
ID ABV78217 standard; DNA; 1653 BP.
XX
AC ABV78217;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human MMP12 DNA SEQ ID NO 101.
XX
KW RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic;
KW virucide; protozoacide; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200255693-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-EP000152.
XX
PR 09-JAN-2001; 2001DE-01000586.
PR 26-OCT-2001; 2001DE-01055280.
PR 29-NOV-2001; 2001DE-01058411.
PR 07-DEC-2001; 2001DE-01060151.
XX
PA (RIBO-) RIBOPHARMA AG.
XX

PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-590671/63.
XX
PT Inhibiting expression of target gene, useful e.g. for inhibiting
PT oncogenes, by administering double-stranded RNA complementary to the
PT target and having an overhang.
XX
PS Claim 10; Page 173; 203pp; German.
XX
CC The invention relates to inhibiting expression of a target gene (I) in a
CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
CC structure of at most 49 consecutive bases. At least part of one strand
CC (as1) of dsRNA1 is complementary to (I) and at least one end of dsRNA1
CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC in humans, also genes in Plasmodium or in viruses or viroids that are
CC pathogenic for humans, animals or plants. Introducing an overhang into
CC dsRNA greatly increases effectiveness for inhibiting gene expression,
CC both in vivo and in vitro and also increases stability and thus the
CC effective concentration inside the cell. The present sequence is that of
CC a gene related to the invention
XX
SQ Sequence 1653 BP; 410 A; 288 C; 280 G; 374 T; 0 U; 301 Other;

Query Match 100.0%; Score 19; DB 6; Length 1653;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||
Db 361 AATTATTGATTCTGTAGGT 343

RESULT 10
ABZ35793/c
ID ABZ35793 standard; DNA; 1653 BP.
XX
AC ABZ35793;
XX
DT 07-FEB-2003 (first entry)
XX
DE Human MMP12 polynucleotide SEQ ID NO 101.
XX
KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW Hepatitis C virus; human papilloma virus; gene; ds.
XX
OS Homo sapiens.
XX
PN DE10100588-A1.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2001; 2001DE-01000588.
XX
PR 09-JAN-2001; 2001DE-01000588.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-683450/74.
XX
PT Inhibiting expression of target genes, useful e.g. for treating tumors,
PT by introducing into cells two double-stranded RNAs that are complementary
PT to the target.
XX
PS Claim 13; Page 70; 100pp; German.
XX
CC The invention relates to inhibiting expression of a target gene in a cell
CC by introducing at least two oligoribonucleotides (dsRNAI and II), both
CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
CC pairs. At least part of one strand (S1, S2) of the ds structures in each
CC of dsRNAI and II are complementary to regions in the target gene. The
CC method uses antisense inhibition of gene expression using double stranded
CC RNA inhibition (RNAi). The method is particularly used to treat tumours
CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
CC humans, animals or plants). The method provides more effective inhibition
CC of expression than known methods using a single dsRNA, even at very low
CC concentrations. When dsRNA has at least one unpaired nucleotide at the
CC end, stability (and thus effective concentration in the cell) is improved
CC and efficiency can be increased further by pretreating the cells with
CC interferon. The present sequence is that of a target DNA of the invention
XX
SQ Sequence 1653 BP; 410 A; 288 C; 280 G; 374 T; 0 U; 301 Other;

Query Match 100.0%; Score 19; DB 6; Length 1653;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||
Db 361 AATTATTGATTCTGTAGGT 343

RESULT 11
ABX10036/c
ID ABX10036 standard; DNA; 1653 BP.
XX
AC ABX10036;
XX
DT 23-JAN-2003 (first entry)
XX
DE Human MMP12 DNA fragment SEQ ID 101.
XX
KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW prion; inhibition; human; ds.
XX
OS Homo sapiens.
XX
PN DE10100587-C1.
XX
PD 21-NOV-2002.
XX
PF 09-JAN-2001; 2001DE-01000587.
XX
PR 09-JAN-2001; 2001DE-01000587.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR WPI; 2002-742209/81.
XX
PT Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT introduction of complementary double-stranded oligoribonucleotide, after
PT treating the cell with interferon.
XX
PS Disclosure; Page 75; 98pp; German.
XX
CC This invention describes a novel method for inhibiting expression of a
CC target gene by introducing into the cell that contains the target gene at
CC least one oligoribonucleotide (dsRNAI) that has a double-stranded (ds)
CC structure of not more than 49 consecutive nucleotides (nt), where at
CC least a segment of one strand of the ds structure is complementary with
CC the target gene and the cells are treated with interferon before
CC introduction of dsRNAI. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX09936-ABX10075 represent

```
CC   gene fragments used to illustrate the method of the invention
XX
SQ   Sequence 1653 BP; 410 A; 288 C; 280 G; 374 T; 0 U; 301 Other;

  Query Match             100.0%;  Score 19;  DB 6;  Length 1653;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 12
ABL91758/c
ID   ABL91758 standard; DNA; 1653 BP.
XX
AC   ABL91758;
XX
DT   28-MAY-2002  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 101.
XX
KW   Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW   Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW   cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100586-C1.
XX
PD   11-APR-2002.
XX
PF   09-JAN-2001; 2001DE-01000586.
XX
PR   09-JAN-2001; 2001DE-01000586.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R,  Limmer S,  Rost S,  Hadwiger P;
XX
DR   WPI; 2002-270454/32.
XX
PT   Inhibiting gene expression in cells, useful for e.g. treating tumors, by
PT   introducing double-stranded complementary oligoRNA having unpaired
PT   terminal bases.
XX
PS   Claim 13; Page 74-75; 104pp; German.
XX
CC   The invention relates to a method for inhibiting expression of a target
CC   gene (ABL91658-ABL91797) in a cell by introducing at least one
CC   oligoribonucleotide that has a double-stranded structure consisting of at
CC   most 49 sequential nucleotide pairs, with at least part of one strand
CC   complementary with the target gene and has at least one end a single-
CC   stranded segment of 1-4 nt. The method provides oligoribonucleotides for
CC   antisense inhibition of gene expression useful e.g. for treating tumours
CC   but the oligoribonucleotides may also be directed against genes present
CC   in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans,
CC   animals or plants) or against cytokine, Id, developmental or prion genes.
CC   The method provides more effective inhibition of gene expression than use
CC   of known oligonucleotides, probably because the unpaired overhang
CC   increases stability and thus intracellular concentration
XX
SQ   Sequence 1653 BP; 410 A; 288 C; 280 G; 374 T; 0 U; 301 Other;

  Query Match             100.0%;  Score 19;  DB 6;  Length 1653;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 13
AAA35052/c
ID   AAA35052 standard; DNA; 1778 BP.
XX
AC   AAA35052;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human adenosine receptor related polynucleotide SEQ ID NO:2741.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US017712.
XX
PR   03-AUG-1998;   98US-0095212P.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstruction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers.
XX
PS   Disclosure; Page 1016-1017; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an antisense
CC   oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC   nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC   impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC   sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC   from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
CC   AAA33992) are specifically claimed ONs from the present invention. N.B.
CC   Sequences given in the disclosure of the present invention do not match
CC   up with their corresponding SEQ ID NO: sequences given in the sequence
CC   listing
XX
SQ   Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 3;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
```

```
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 14
AAF21174/c
ID   AAF21174 standard; DNA; 1778 BP.
XX
AC   AAF21174;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Human low adenosine antisense oligonucleotide related sequence #2741.
XX
KW   Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW   human; airway disorder; bronchoconstriction; lung inflammation;
KW   surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW   immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW   respiratory obstruction; pulmonary obstruction; impeded respiration;
KW   surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW   respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW   pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW   chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200062736-A2.
XX
PD   26-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US008020.
XX
PR   06-APR-1999;   99US-0127958P.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
PA   (NYCE/) NYCE J W.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-679539/66.
XX
PT   Low adenosine (A) content antisense oligonucleotides which do not trigger
PT   adenosine receptors during metabolism, useful e.g. for treating cancers
PT   and respiratory obstructions.
XX
PS   Disclosure; Page 1094-1095; 1592pp; English.
XX
CC   The present invention describes low adenosine (A) content antisense
CC   oligonucleotides and compositions (I) comprising them. In the antisense
CC   oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC   (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC   immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC   The antisense oligonucleotides and (I) can be used to down-regulate the
CC   expression and or activity of target polypeptides associated with
CC   lung/respiratory disorders and malignancies, such as stimulating and
CC   activating peptide factors and transmitters, transcription factors,
CC   immunoglobulins and antibodies, antibody receptors, cytokines and
CC   chemokines, endogenously produced specific and non-specific enzymes,
CC   binding proteins, adhesion molecules and their receptors, cytokine and
CC   chemokine receptors, adenosine receptors, bradykinin receptors, central
CC   nervous system (CNS) and peripheral nervous and non-nervous system
CC   receptors, CNS and peripheral nervous and non-nervous system peptide
CC   transmitters, defensins, growth factors, vasoactive peptides and
CC   receptors, binding proteins and malignancy associated proteins. The
CC   antisense oligonucleotides may be used in this way to treat disorders
CC   including respiratory obstruction (especially pulmonary obstruction
CC   and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC   surfactant hypoproduction which are associated with a disease or
CC   condition selected from pulmonary vasoconstriction, inflammation,
CC   allergies, asthma, impeded respiration, respiratory distress syndrome
CC   (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC   pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC   and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC   fragments and antisense oligonucleotides used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 3;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 15
AAH28229/c
ID   AAH28229 standard; cDNA; 1778 BP.
XX
AC   AAH28229;
XX
DT   05-SEP-2001  (first entry)
XX
DE   Nucleotide sequence of matrix metalloproteinase-12.
XX
KW   Growth factor; protein inhibitor; protease; damaged tissue;
KW   platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
KW   connective tissue derived growth factor; CTGF; chrysalin; VEGF;
KW   keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
KW   transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
KW   granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
KW   vascular endothelial growth factor; urokinase plasminogen activator;
KW   dermal ulcer; wound; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             13. .1425
FT                   /*tag=  a
FT                   /product= "MMP-12"
XX
PN   WO200149309-A2.
XX
PD   12-JUL-2001.
XX
PF   21-DEC-2000; 2000WO-IB001935.
XX
PR   29-DEC-1999;   99GB-00030768.
XX
PA   (PFIZ ) PFIZER LTD.
PA   (PFIZ ) PFIZER INC.
XX
PI   Davies MJ,  Huggins JP,  Mcintosh FS,  Occleston NL;
XX
DR   WPI; 2001-418351/44.
DR   P-PSDB; AAB84614.
XX
PT   Composition for the treatment of damaged tissue i.e. chronic wounds and
PT   dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
PT   factor.
XX
PS   Disclosure; Page 560-561; 572pp; English.
XX
CC   The specification describes a pharmaceutical composition, comprising a
CC   growth factor, an inhibitor agent, i.e. a protease. The inhibitor agent
CC   inhibits the action of at least one specific adverse protein, i.e. a
CC   protease, that is upregulated in a damaged tissue such as a wound
CC   environment. Growth factors which are included in the composition of the
CC   invention are platelet-derived growth factor (PDGF), fibroblast growth
```

CC factor (FGF), connective tissue derived growth factor (CTGF),
CC keratinocyte-derived growth factor (KGF), transforming growth factor-beta
CC (TGF-beta), granulocyte macrophage colony stimulating factor (GM-CSF),
CC epidermal growth factor (EGF), vascular endothelial growth factor (VEGF),
CC and chrysalin. Inhibitors which are included in the composition of the
CC invention include inhibitors of urokinase-type plasminogen activator
CC (uPA) and matrix metalloproteinase (MMP). The composition is useful for
CC the treatment of chronic damaged tissue, i.e. wounds and dermal ulcers.
CC The present sequence encodes a human MMP-12, and is used to produce the
CC composition of the invention
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 4; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 16
ABL62664/c
ID ABL62664 standard; DNA; 1778 BP.
XX
AC ABL62664;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:1001.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 1001; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 17
ABL65817/c
ID ABL65817 standard; DNA; 1778 BP.
XX
AC ABL65817;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4154.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.

XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 4154; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
|||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 18
ABL64743/c
ID ABL64743 standard; DNA; 1778 BP.
XX
AC ABL64743;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3080.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.

PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 3080; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 19
ABL62102/c
ID ABL62102 standard; DNA; 1778 BP.
XX
AC ABL62102;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:439.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 439; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,

```
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 20
ABL66489/c
ID   ABL66489 standard; DNA; 1778 BP.
XX
AC   ABL66489;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Lung cancer related gene sequence SEQ ID NO:4826.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US010838.
XX
PR   05-JUN-2000; 2000US-0209473P.
PR   05-JUN-2000; 2000US-0209531P.
PR   18-SEP-2000; 2000US-0233133P.
PR   18-SEP-2000; 2000US-0233617P.
PR   20-SEP-2000; 2000US-0234009P.
PR   20-SEP-2000; 2000US-0234034P.
PR   20-SEP-2000; 2000US-0234052P.
PR   22-SEP-2000; 2000US-0234509P.
PR   22-SEP-2000; 2000US-0234567P.
PR   25-SEP-2000; 2000US-0234923P.
PR   25-SEP-2000; 2000US-0234924P.
PR   25-SEP-2000; 2000US-0235077P.
PR   25-SEP-2000; 2000US-0235082P.
PR   25-SEP-2000; 2000US-0235134P.
PR   25-SEP-2000; 2000US-0235280P.
PR   26-SEP-2000; 2000US-0235637P.
PR   26-SEP-2000; 2000US-0235638P.
PR   27-SEP-2000; 2000US-0235711P.
PR   27-SEP-2000; 2000US-0235720P.
PR   27-SEP-2000; 2000US-0235840P.
PR   27-SEP-2000; 2000US-0235863P.
PR   28-SEP-2000; 2000US-0236028P.
PR   28-SEP-2000; 2000US-0236032P.
PR   28-SEP-2000; 2000US-0236033P.
PR   28-SEP-2000; 2000US-0236034P.
PR   28-SEP-2000; 2000US-0236109P.
PR   28-SEP-2000; 2000US-0236111P.
PR   29-SEP-2000; 2000US-0236842P.
PR   29-SEP-2000; 2000US-0236891P.
PR   02-OCT-2000; 2000US-0237172P.
PR   02-OCT-2000; 2000US-0237173P.
PR   02-OCT-2000; 2000US-0237278P.
PR   02-OCT-2000; 2000US-0237294P.
PR   02-OCT-2000; 2000US-0237295P.
PR   02-OCT-2000; 2000US-0237316P.
PR   03-OCT-2000; 2000US-0237425P.
PR   03-OCT-2000; 2000US-0237598P.
PR   03-OCT-2000; 2000US-0237604P.
PR   03-OCT-2000; 2000US-0237606P.
PR   03-OCT-2000; 2000US-0237608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI   Soppet DR,  Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 4826; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 21
ABN95681/c
ID   ABN95681 standard; DNA; 1778 BP.
XX
AC   ABN95681;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #2179 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030589.
XX
```

PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 2179; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 22
ABK54023/c
ID ABK54023 standard; cDNA; 1778 BP.
XX
AC ABK54023;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human head and neck tumour cDNA, SEQ ID No 218.
XX
KW Human; head and neck cancer; tumour; cytostatic; immunogenic; vaccine;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200212329-A2.
XX
PD 14-FEB-2002.
XX
PF 01-AUG-2001; 2001WO-US024226.
XX
PR 03-AUG-2000; 2000US-0223281P.
PR 16-NOV-2000; 2000US-0249933P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Fan L;
XX
DR WPI; 2002-257467/30.
DR P-PSDB; AAU91061.
XX
PT Novel polynucleotide encoding head and neck tumor polypeptides, useful in
PT pharmaceutical compositions, e.g. vaccines, for treating head and neck
PT cancers.
XX
PS Claim 1; Page 180-181; 200pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising
CC sequences selected from 273 sequences fully defined in the specification.
CC (I), including its encoded polypeptide (II), an antibody binding to (II),
CC a fusion protein comprising (II) and a T-cell population stimulated by
CC (I) or (II) are useful for stimulating an immune response in a patient
CC and treating head and neck cancer in a patient. An oligonucleotide (III)
CC that hybridises to (I) is useful for determining the presence of cancer
CC in a patient, by obtaining a biological sample from the patient,
CC contacting the sample with (III), detecting in the sample an amount of a
CC polynucleotide that hybridises to the oligonucleotide, and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotides to a
CC predetermined cut-off value. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis and
CC treatment of head and neck cancer. ABK53806-ABK54078 represent human head
CC and neck cancer cDNA sequences of the invention
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 23
ACC51017/c
ID ACC51017 standard; cDNA; 1778 BP.
XX
AC ACC51017;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated cDNA sequence SEQ ID NO:124.
XX
KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003003906-A2.
XX
PD 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US021338.
XX
PR 03-JUL-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Aziz N;
XX
DR WPI; 2003-201532/19.
DR P-PSDB; ABR48203.
XX
PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
PS Claim 6; Page 274-275; 307pp; English.
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises

CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridises to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 1778 BP; 549 A; 357 C; 337 G; 535 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 8; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 24
ABX76137/c
ID ABX76137 standard; DNA; 1778 BP.
XX
AC ABX76137;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polynucleotide #14.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
PN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
DR P-PSDB; ABU56421.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 22; Page 198; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention
XX
SQ Sequence 1778 BP; 549 A; 357 C; 337 G; 535 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 8; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 25
ACF12895/c
ID ACF12895 standard; cDNA; 1778 BP.
XX
AC ACF12895;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human cervical cancer cell marker protein SEQ ID NO:134.
XX
KW Human; cervical cancer; cervical cancer marker; cancer therapy;
KW detection; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2002101075-A2.
XX
PD 19-DEC-2002.
XX
PF 12-JUN-2002; 2002WO-US018638.
XX
PR 13-JUN-2001; 2001US-0298155P.
PR 13-JUN-2001; 2001US-0298159P.
PR 14-NOV-2001; 2001US-0335936P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
PI Gannavarapu M, Glatt K, Hoersch S;
XX
DR WPI; 2003-156967/15.
DR P-PSDB; ABR92113.
XX
PT New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
PS Claim 4; Page 288; 386pp; English.
XX
CC ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (M1) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's

CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 8; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 26
ADD18687/c
ID ADD18687 standard; DNA; 1778 BP.
XX
AC ADD18687;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID118.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnerary; gene therapy;
KW hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003018621-A2.
XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
DR WPI; 2003-290046/28.
DR P-PSDB; ADD18686.
XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 118; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumourigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 27
ADF18049/c
ID ADF18049 standard; cDNA; 1778 BP.
XX
AC ADF18049;
XX
DT 12-FEB-2004 (first entry)
XX
DE cDNA encoding human matrix metalloproteinase 12 (MMP-12).
XX
KW anoretic; antidiabetic; immunomodulator; MMP-12 agonist;
KW MMP-12 antagonist; MMP-12 inverse agonist; anti-MMP-12 antibody;
KW antisense MMP-12; MMP-12 ribozyme; lipid homeostasis modulator;
KW glucose homeostasis modulator; insulin homeostasis modulator;
KW adipocyte growth modulator; metabolic disorder;
KW matrix metalloproteinase 12; MMP-12; translation inhibition; obesity;
KW overweight; diabetes; insulin resistance; cachexia; anorexia; gene; ss;
KW human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 13. .1425
FT /*tag= a
FT /product= "Human MMP-12"
FT /note= "Matrix metalloproteinase"
XX
PN US2003157110-A1.
XX
PD 21-AUG-2003.
XX
PF 07-JAN-2003; 2003US-00337632.
XX
PR 07-JAN-2002; 2002US-0346354P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI An WF, Chen H;
XX
DR WPI; 2003-897814/82.
DR P-PSDB; ADF18050.
XX
PT Identifying a compound capable of treating a metabolic disorder comprises
PT assaying for compounds that modulate matrix metalloprotease MMP-12
PT expression or activity.
XX
PS Claim 4; SEQ ID NO 1; 38pp; English.
XX
CC The invention describes a method of identifying a compound capable of
CC treating a metabolic disorder comprising assaying the ability of the
CC compound to modulate a matrix metalloproteinase 12 (MMP-12) nucleic acid
CC expression or MMP-12 polypeptide activity. The compound has anoretic,
CC antidiabetic and immunomodulator properties. An antisense oligonucleotide
CC or ribozyme targets the MMP-12 transcripts and inhibits translation.
CC Alternatrively the oligonucleotide forms a triple helix with the promoter

CC of the MMP-12 gene and inhibits transcription. The invention is used to
CC identify a compound useful for treating a metabolic disorder,
CC particularly obesity, overweight, diabetes, insulin resistance, cachexia
CC or anorexia. This sequence encodes human matrix metalloproteinase 12 (MMP
CC -12).
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 28
ADF74444/c
ID ADF74444 standard; DNA; 1778 BP.
XX
AC ADF74444;
XX
DT 26-FEB-2004 (first entry)
XX
DE Full length human matrix metalloproteinase-12 (MMP-12) DNA (SeqID 14).
XX
KW gene; ds; human; microbial infection; MMPAP-12; MMP-12;
KW matrix metalloproteinase-12; matrix degrading enzyme; emphysema;
KW aortic aneurysm; antimicrobial; pneumonia; peritonitis; antibacterial;
KW virucidal; protozoacidal; fungicidal.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 13. .1425
FT /*tag= a
FT /product= "MMP-12 protein"
XX
PN WO2003087325-A2.
XX
PD 23-OCT-2003.
XX
PF 08-APR-2003; 2003WO-US010911.
XX
PR 08-APR-2002; 2002US-0370649P.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Shapiro SD, Hartzell WO;
XX
DR WPI; 2003-865359/80.
DR P-PSDB; ADF74443.
XX
PT New isolated MMPAP-12 polypeptide and nucleic acid molecule, useful for
PT treating infectious diseases associated with microorganisms, bacteria,
PT virus, fungi or parasites.
XX
PS Claim 18; SEQ ID NO 14; 126pp; English.
XX
CC This invention relates to a novel method for treating or preventing
CC microbial infections by the administration of MMPAP-12 nucleic acids and
CC encoded polypeptides thereof. Specifically MMPAP-12, also referred to as
CC MMP-12, refers to matrix metalloproteinase-12 and is a member of the
CC family of matrix degrading enzymes capable of matrix destruction and
CC hence contributes to such diseases as emphysema and aortic aneurysm. The
CC present invention, however, describes the unexpected non-enzymatic,
CC antimicrobial properties of the MMP-12 protein at the cellular level. As
CC such, MMP-12 nucleic acids, proteins and fragments thereof, have been
CC identified that exhibit specific activities which are useful for the
CC treatment of infectious diseases associated with microorganisms,
CC bacteria, virus, fungi or parasites such as pneumonia or peritonitis.
CC Accordingly, these compositions have been described as antibacterial,
CC virucidal, protozoacidal and fungicidal. This polynucleotide is the full
CC length human MMP-12 DNA sequence of the invention.
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 29
ACC46771/c
ID ACC46771 standard; cDNA; 1778 BP.
XX
AC ACC46771;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human COPD related protein encoding cDNA SEQ ID NO:22.
XX
KW Human; chronic obstructive pulmonary disease; COPD; chronic lung disease;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200297127-A2.
XX
PD 05-DEC-2002.
XX
PF 28-MAY-2002; 2002WO-EP005835.
XX
PR 31-MAY-2001; 2001GB-00013266.
XX
PA (FARB) BAYER AG.
XX
PI Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;
XX
DR WPI; 2003-140492/13.
DR P-PSDB; ABP96800.
XX
PT Predicting, diagnosing or prognosing chronic lung disease, by detecting a
PT chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX
PS Claim 8; Page 110-111; 214pp; English.
XX
CC The present invention describes a method for predicting, diagnosing or
CC prognosing chronic lung disease by detecting a chronic obstructive
CC pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
CC ACC46777, which encode the COPD related proteins in ABP96779 to
CC ABP96806). The method is useful for predicting, diagnosing or prognosing
CC chronic lung disease in a biological sample. The COPD genes and proteins
CC encoded by them from the present invention (I) can be used for treating
CC or preventing chronic lung disease in a mammal. (I) can be used in an
CC animal model for determining the efficacy, toxicity, or side effects of
CC treatment with (I), and determining the mechanism of action of (I).
CC ACC46778 to ACC46903 represent COPD related PCR primers and probes used
CC in an example from the present invention
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 30
ABZ96868/c
ID ABZ96868 standard; DNA; 1778 BP.
XX
AC ABZ96868;
XX
DT 17-OCT-2003 (first entry)
XX
DE Human nucleic acid sequence.
XX
KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
XX
OS Homo sapiens.
XX
PN WO200285308-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013135.
XX
PR 24-APR-2001; 2001US-0286137P.
XX
PA (EPIG-) EPIGENESIS PHARM INC.
XX
PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX
DR WPI; 2003-229219/22.
XX
PT Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiquinone.
XX
PS Disclosure; SEQ ID NO 12110; 872pp; English.
XX
CC The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antiinflammatory steroid and ubiquinone. A composition of the invention
CC has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
CC immunosuppressive, and cytostatic activity. The composition may have a
CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory, lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiquinone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 31
ABD20717/c
ID ABD20717 standard; DNA; 1778 BP.
XX
AC ABD20717;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human pulmonary and inflammatory target DNA #328.
XX
KW Human; antisense; bronchoconstriction; allergy; hyposecretion; pain;
KW respiratory tract inflammation; adenosine sensitivity; lung; cancer;
KW surfactant depletion; antiallergic; antiinflammatory; antiasthmatic;
KW analgesic; hypotensive; immunosuppressive; cytostatic; cystic fibrosis;
KW beta-adrenergic agonist; respiratory disease; pulmonary vasoconstriction;
KW respiratory distress syndrome; allergic rhinitis; pulmonary hypertension;
KW emphysema; chronic obstructive pulmonary disease; cancer; bronchitis;
KW pulmonary transplantation rejection; ds.
XX
OS Homo sapiens.
XX
PN WO200285309-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013143.
XX
PR 24-APR-2001; 2001US-0286036P.
XX
PA (EPIG-) EPIGENESIS PHARM INC.
XX
PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX
DR WPI; 2003-093058/08.
XX
PT Pharmaceutical composition for treating asthma, has antisense
PT oligonucleotide containing less percentage of adenosine, targeted to
PT nucleic acids associated with lung airway or lung dysfunction, and
PT bronchodilating agent.
XX
PS Claim 15; SEQ ID NO 12110; 763pp; English.
XX
CC This invention describes a novel composition (a) a first active agent,
CC comprising oligonucleotides, effective for alleviating
CC bronchoconstriction, respiratory tract inflammation, allergies and
CC reducing adenosine sensitivity, levels of adenosine (A) or (A) receptors,
CC surfactant depletion or hyposecretion, when administered to a mammal. The
CC oligonucleotides are derived from a gene encoding or regulating
CC expression of a target polypeptide associated with lung airway or lung
CC dysfunction or cancer and can be anti-sense to the corresponding mRNA.
CC The invention also describes a kit, that comprises: (a) a delivery
CC device, in separate containers, (b) the oligonucleotides, (c)
CC instructions for adding a carrier and for use of the kit. The composition
CC of the invention has antiallergic, antiinflammatory, antiasthmatic,
CC analgesic, hypotensive, immunosuppressive and cytostatic activity, is a
CC beta-adrenergic agonist. The composition is useful for preventing or
CC treating a respiratory, lung or malignant disease. The administered
CC composition comprises oligo and is administered to reduce the production
CC or availability, or to increase the degradation of the target mRNA or to
CC reduce the amount of target polypeptide present in the lungs. The
CC pulmonary obstruction, and/or bronchoconstriction and/or lung
CC inflammation, allergies and/or surfactant hypoproduction are associated
CC with a disease or condition such as pulmonary vasoconstriction,
CC inflammation, allergies, asthma, impeded respiration, respiratory
CC distress syndrome, pain, cystic fibrosis, allergic rhinitis, pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease, pulmonary
CC transplantation rejection, pulmonary infections, bronchitis or cancer.
CC The reduced adenosine content of the anti-sense oligos corresponding to
CC thymidines present in the target RNA serves to prevent the breakdown of
CC the oligonucleotides into products that free adenosine into the system
CC e.g., lung, brain, heart, kidney, etc, tissue environment and thereby, to
CC prevent any unwanted effects due to it

XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 11; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 32
ADI53688/c
ID ADI53688 standard; cDNA; 1778 BP.
XX
AC ADI53688;
XX
DT 22-APR-2004 (first entry)
XX
DE Human matrix metalloproteinase 12, MMP-12, cDNA sequence, SEQ ID 1.
XX
KW Antiinflammatory; Antipsoriatic; Antiasthmatic; Antiarthritic;
KW Respiratory; antisense oligonucleotide; matrix metalloproteinase 12;
KW MMP-12; inflammatory bowel disease; ulcerative colitis; Crohn's disease;
KW rheumatoid arthritis; psoriasis; emphysema; asthma; human; ss.
XX
OS Homo sapiens.
XX
PN WO2004009098-A1.
XX
PD 29-JAN-2004.
XX
PF 17-JUL-2003; 2003WO-SE001223.
XX
PR 18-JUL-2002; 2002SE-00002253.
PR 04-SEP-2002; 2002US-0407680P.
XX
PA (INDE-) INDEX PHARM AB.
XX
PI Dieckmann A, Loefberg R, Von Stein O, Von Stein P, Good L;
XX
DR WPI; 2004-123288/12.
XX
PT New compound having a sequence targeted to a nucleic acid encoding
PT metalloproteinase 12 (MMP-12), useful for preparing a composition for
PT treating or preventing MMP-12 dependent disorder in a human patient e.g.,
PT asthma or psoriasis.
XX
PS Claim 2; SEQ ID NO 1; 55pp; English.
XX
CC The present invention relates to antisense oligonucleotides (ADI53690-
CC ADI53701) for matrix metalloproteinase 12 (MMP-12; ADI53688 and
CC ADI53689), which specifically hybridise with the nucleic acid encoding
CC MMP-12 and inhibiting the translation of MMP-12 protein. The antisense
CC oligonucleotides are useful for preparing a composition for treating or
CC preventing MMP-12 dependent disorder in a human patient e.g. inflammatory
CC bowel disease, such as ulcerative colitis and Crohn's disease, rheumatoid
CC arthritis, psoriasis, emphysema or asthma.
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 33
ADN05808/c
ID ADN05808 standard; cDNA; 1778 BP.
XX
AC ADN05808;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #1135.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR P-PSDB; ADN05809.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 1; SEQ ID NO 2202; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 34
ADQ29584/c
ID ADQ29584 standard; DNA; 1778 BP.
XX
AC ADQ29584;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human colorectal cancer-associated protein coding sequence #6.
XX
KW human; colon cancer; TIMP1; Reg1-alpha;
KW colorectal cancer-associated marker; gene; ds.
XX
OS Homo sapiens.
XX
PN EP1439393-A2.
XX
PD 21-JUL-2004.
XX
PF 15-DEC-2003; 2003EP-00257868.
XX
PR 13-DEC-2002; 2002US-0433554P.

PR 31-JUL-2003; 2003US-0491397P.
XX
PA (FARB) BAYER HEALTHCARE LLC.
PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.
XX
PI Astle JH, Boardman LA, Bugart LJ, Burgess CC, Catino TJ;
PI Dwivedi P, Huntress M, Johnson KA, Lewis ME, Maimonis PJ, Myerow SH;
PI Brown-Shimer SLA, Thiagalingam A, Thibodeau SN, Molino GA;
XX
DR WPI; 2004-545561/53.
DR P-PSDB; ADQ29651.
XX
PT Diagnosing colon cancer in individual, preferably human, by detecting
PT presence of TIMP 1 in sample, where presence of TIMP 1 in sample is
PT indicative of colon cancer in individual.
XX
PS Claim 7; SEQ ID NO 10; 433pp; English.
XX
CC The invention comprises a method for diagnosing colon cancer in an
CC individual, the method involves obtaining a serum sample from the
CC individual and detecting the presence of either TIMP1 or Reg1-alpha and
CC an additional colorectal cancer-associated marker. The method of the
CC invention is useful for diagnosing colon cancer in an individual. The
CC present DNA sequence represents a human colorectal cancer-associated
CC protein coding sequence of the invention.
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 35
ADR24908/c
ID ADR24908 standard; DNA; 1778 BP.
XX
AC ADR24908;
XX
DT 21-OCT-2004 (first entry)
XX
DE Breast cancer prognosis marker #769.
XX
KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004065545-A2.
XX
PD 05-AUG-2004.
XX
PF 15-JAN-2004; 2004WO-US001100.
XX
PR 15-JAN-2003; 2003US-00342887.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
PA (NECA-) NETHERLANDS CANCER INST.
XX
PI Van't Veer LJ, He Y;
XX
DR WPI; 2004-593473/57.
XX
PT Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
PS Disclosure; SEQ ID NO 769; 226pp; English.
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
SQ Sequence 1778 BP; 548 A; 357 C; 337 G; 536 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 13; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 36
ABZ20476/c
ID ABZ20476 standard; cDNA; 1781 BP.
XX
AC ABZ20476;
XX
DT 23-JAN-2003 (first entry)
XX
DE Matrix metalloproteinase 12 full length cDNA sequence SEQ ID NO:2910.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US010421.
XX
PR 30-MAR-2001; 2001US-0280255P.
PR 28-AUG-2001; 2001US-0315563P.
PR 09-JAN-2002; 2002US-0347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
DR P-PSDB; ABP54460.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT virology, immunology, microbiology, molecular biology and recombinant DNA
PT techniques.
XX
PS Claim 1; SEQ ID NO 2910; 207pp; English.
XX
CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytostatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques. N.B. The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

```
SQ   Sequence 1781 BP; 547 A; 359 C; 341 G; 534 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 8;  Length 1781;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 37
AAF81624/c
ID   AAF81624 standard; cDNA; 1818 BP.
XX
AC   AAF81624;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Human macrophage metalloelastase HME coding sequence.
XX
KW   Human; HME: macrophage metalloelastase; elastolysis; pulmonary emphysema;
KW   inflammatory disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1413
FT                   /*tag=  a
FT                   /product= "macrophage metalloelastase"
XX
PN   US6204043-B1.
XX
PD   20-MAR-2001.
XX
PF   01-MAR-1995;   95US-00396988.
XX
PR   28-MAY-1993;   93US-00068392.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Shapiro SD;
XX
DR   WPI; 2001-289518/30.
DR   P-PSDB; AAB74595.
XX
PT   New human macrophage metalloelastase, useful in the normal embryonic
PT   development, growth, tissue remodeling and tissue repair, particularly in
PT   studying the pathogenesis of pulmonary emphysema.
XX
PS   Example; Fig 5; 20pp; English.
XX
CC   The present invention provides the protein and coding sequences of human
CC   macrophage metalloelastase (HME). The protein has elastolytic activity.
CC   HME has a role in tissue remodeling and repair associated with
CC   development and inflammation, and abnormal expression can result in
CC   tumour invasiveness, arthritis and atherosclerosis. It is also thought to
CC   contribute to the pathogenesis of pulmonary emphysema and other
CC   inflammatory destructive diseases. The present sequence is the HME coding
CC   sequence
XX
SQ   Sequence 1818 BP; 563 A; 365 C; 350 G; 540 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 1818;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 38
AAC64994/c
ID   AAC64994 standard; cDNA; 1818 BP.
XX
AC   AAC64994;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human macrophage metalloelastase coding sequence.
XX
KW   Human; macrophage metalloproteinase; elastin; matrix degrading enzyme;
KW   emphysema; ss.
XX
OS   Homo sapiens.
XX
PN   US6150152-A.
XX
PD   21-NOV-2000.
XX
PF   28-MAY-1993;   93US-00068392.
XX
PR   28-MAY-1993;   93US-00068392.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Shapiro SD;
XX
DR   WPI; 2001-049090/06.
DR   P-PSDB; AAB49982.
XX
PT   Novel human macrophage metalloelastase polynucleotides and polypeptides
PT   useful for measuring elastin degradation.
XX
PS   Claim 2; Fig 5; 20pp; English.
XX
CC   The present invention provides the coding and protein sequences for the
CC   human macrophage metalloelastase. This is a matrix degrading
CC   metalloproteinase which has the ability to degrade elastin, and can be
CC   used to measure elastin degradation
XX
SQ   Sequence 1818 BP; 563 A; 365 C; 350 G; 540 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 1818;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 39
ADQ23366/c
ID   ADQ23366 standard; DNA; 1873 BP.
XX
AC   ADQ23366;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 6186.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
```

PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 6186; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 1873 BP; 590 A; 379 C; 356 G; 548 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 1873;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 421 AATTATTGATTCTGTAGGT 403

RESULT 40
AAC77981/c
ID AAC77981 standard; cDNA; 1874 BP.
XX
AC AAC77981;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated gene sequence SEQ ID NO:375.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR P-PSDB; AAB43772.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
XX
PS Claim 1; Page 921-922; 2352pp; English.
XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerary; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 1874 BP; 609 A; 367 C; 352 G; 546 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 3; Length 1874;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 423 AATTATTGATTCTGTAGGT 405

RESULT 41
ADB47336/c
ID ADB47336 standard; cDNA; 1988 BP.
XX
AC ADB47336;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA upregulated in dendritic cells SEQ ID NO 36.
XX
KW ss; gene; human; dendritic cells; high throughput; cancer;
KW infectious disease; autoimmune disease; allergy;
KW graft versus host disease; vaccine enhancing; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003134283-A1.
XX
PD 17-JUL-2003.
XX
PF 03-OCT-2001; 2001US-00971392.
XX
PR 03-OCT-2000; 2000US-0237652P.
XX
PA (PETE/) PETERSON D P.
PA (PEAR/) PEARSON C I.
PA (COCK/) COCKS B G.
XX
PI Peterson DP, Pearson CI, Cocks BG;
XX
DR WPI; 2003-662509/62.
XX

PT New combination comprises cDNAs that are differentially expressed in
PT dendritic cells useful for preparing a composition for diagnosing or
PT treating cancer, infectious disease, autoimmunity, allergy or graft
PT versus host disease.
XX
PS Claim 1; SEQ ID NO 36; 28pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in dendritic cells (DC). Also included is a high
CC throughput method for detecting differential expression of one or more
CC cDNAs in a sample containing nucleic acids. The combination is useful for
CC preparing a composition for diagnosing, treating and monitoring the
CC treatment of cancer, infectious disease, autoimmunity, allergy or graft
CC versus host disease, or for enhancing a vaccine. The present sequence
CC represents a human cDNA upregulated in dendritic cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=20030134283.
XX
SQ Sequence 1988 BP; 627 A; 387 C; 371 G; 602 T; 0 U; 1 Other;

Query Match 100.0%; Score 19; DB 10; Length 1988;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 396 AATTATTGATTCTGTAGGT 378

RESULT 42
ADD29857/c
ID ADD29857 standard; mRNA; 2870 BP.
XX
AC ADD29857;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human tumour suppressor mRNA SEQ ID NO:288.
XX
KW ss; human; tumour suppressor; cancer; cancer; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2003058201-A2.
XX
PD 17-JUL-2003.
XX
PF 31-DEC-2002; 2002WO-US041825.
XX
PR 31-DEC-2001; 2001US-0345317P.
XX
PA (QUAR-) QUARK BIOTECH INC.
PA (CLEV-) CLEVELAND CLINIC FOUND.
XX
PI Feinstein E, Gudkov AV;
XX
DR WPI; 2003-598393/56.
XX
PT Diagnosing cancer comprises determining the polypeptide or polynucleotide
PT levels e.g., hepatic lipase, in a sample from a subject, where a higher
PT level compared to that in a subject free of cancer is indicative of
PT cancer.
XX
PS Disclosure; SEQ ID NO 288; 272pp; English.
XX
CC The invention relates to a novel method for diagnosing a cancer in a
CC subject. the method comprises determining, in a sample from the subject,
CC the level of at least one polypeptide, where a higher level of the
CC polypeptide compared to the level of the polypeptide in a subject free of
CC cancer is indicative of cancer. The polypeptide is selected from any of
CC the polypeptides encoded by the polynucleotides listed in the
CC specification and polypeptides which are at least 70% homologous to the
CC polypeptides. The method of the invention has cytostatic activity, and
CC may have a use in gene therapy. The method is useful in identifying
CC markers specific for one or several types of cancer, depending on the
CC tissue origin, which may be used in numerous diagnostic and prognostic
CC applications as well as cancer type-specific targets for therapeutic
CC intervention. The compounds that modulate the activity of a tumour
CC suppressor gene are useful in the treatment of cancer or as anti-cancer
CC drugs. The present sequence represents a polynucleotide of the invention.
XX
SQ Sequence 2870 BP; 883 A; 540 C; 563 G; 884 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 2870;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 1465 AATTATTGATTCTGTAGGT 1447

RESULT 43
AAA35055/c
ID AAA35055 standard; DNA; 9137 BP.
XX
AC AAA35055;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide SEQ ID NO:2744.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US017712.
XX
PR 03-AUG-1998; 98US-0095212P.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstruction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers.
XX
PS Disclosure; Page 1018-1020; 1343pp; English.
XX
CC The present invention describes a new composition comprising an antisense
CC oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive

```
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC   sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC   from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
CC   AAA33992) are specifically claimed ONs from the present invention. N.B.
CC   Sequences given in the disclosure of the present invention do not match
CC   up with their corresponding SEQ ID NO: sequences given in the sequence
CC   listing
XX
SQ   Sequence 9137 BP; 2510 A; 2118 C; 2093 G; 2416 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 3;  Length 9137;
  Best Local Similarity   100.0%;  Pred. No. 0.32;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db       4780 AATTATTGATTCTGTAGGT 4762


RESULT 44
AAF21177/c
ID   AAF21177 standard; DNA; 9137 BP.
XX
AC   AAF21177;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Human low adenosine antisense oligonucleotide related sequence #2744.
XX
KW   Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW   human; airway disorder; bronchoconstriction; lung inflammation;
KW   surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW   immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW   respiratory obstruction; pulmonary obstruction; impeded respiration;
KW   surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW   respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW   pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW   chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200062736-A2.
XX
PD   26-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US008020.
XX
PR   06-APR-1999;   99US-0127958P.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
PA   (NYCE/) NYCE J W.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-679539/66.
XX
PT   Low adenosine (A) content antisense oligonucleotides which do not trigger
PT   adenosine receptors during metabolism, useful e.g. for treating cancers
PT   and respiratory obstructions.
XX
PS   Disclosure; Page 1096-1098; 1592pp; English.
XX
CC   The present invention describes low adenosine (A) content antisense
CC   oligonucleotides and compositions (I) comprising them. In the antisense
CC   oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC   (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC   immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC   The antisense oligonucleotides and (I) can be used to down-regulate the
CC   expression and or activity of target polypeptides associated with
CC   lung/respiratory disorders and malignancies, such as stimulating and
CC   activating peptide factors and transmitters, transcription factors,
CC   immunoglobulins and antibodies, antibody receptors, cytokines and
CC   chemokines, endogenously produced specific and non-specific enzymes,
CC   binding proteins, adhesion molecules and their receptors, cytokine and
CC   chemokine receptors, adenosine receptors, bradykinin receptors, central
CC   nervous system (CNS) and peripheral nervous and non-nervous system
CC   receptors, CNS and peripheral nervous and non-nervous system peptide
CC   transmitters, defensins, growth factors, vasoactive peptides and
CC   receptors, binding proteins and malignancy associated proteins. The
CC   antisense oligonucleotides may be used in this way to treat disorders
CC   including respiratory obstruction (especially pulmonary obstruction
CC   and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC   surfactant hypoproduction which are associated with a disease or
CC   condition selected from pulmonary vasoconstriction, inflammation,
CC   allergies, asthma, impeded respiration, respiratory distress syndrome
CC   (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC   pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC   and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC   fragments and antisense oligonucleotides used in the exemplification of
CC   the present invention
XX
SQ   Sequence 9137 BP; 2510 A; 2118 C; 2093 G; 2416 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 3;  Length 9137;
  Best Local Similarity   100.0%;  Pred. No. 0.32;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db       4780 AATTATTGATTCTGTAGGT 4762


RESULT 45
ABZ96871/c
ID   ABZ96871 standard; DNA; 9137 BP.
XX
AC   ABZ96871;
XX
DT   17-OCT-2003  (first entry)
XX
DE   Human nucleic acid sequence.
XX
KW   Human; antisense; lung dysfunction; nasal airway dysfunction;
KW   antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW   antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW   antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW   adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW   lung inflammation; respiratory disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285308-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013135.
XX
PR   24-APR-2001; 2001US-0286137P.
XX
PA   (EPIG-) EPIGENESIS PHARM INC.
XX
PI   Nyce JW,  Li Y,  Sandrasagra A,  Katz E,  Pabalan J,  Aguilar D;
PI   Miller S,  Tang L,  Shahabuddin S;
XX
DR   WPI; 2003-229219/22.
XX
PT   Pharmaceutical composition for treating ailments associated with impaired
```

PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiquinone.
XX
PS Disclosure; SEQ ID NO 12113; 872pp; English.
XX
CC The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antiinflammatory steroid and ubiquinone. A composition of the invention
CC has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
CC immunosuppressive, and cytostatic activity. The composition may have a
CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory, lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiquinone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9137 BP; 2510 A; 2118 C; 2093 G; 2416 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 9137;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db       4780 AATTATTGATTCTGTAGGT 4762
```

Search completed: April 26, 2005, 15:31:22
Job time : 443 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 26, 2005, 15:31:33 ; Search time 136 Seconds
(without alignments)
228.597 Million cell updates/sec

Title: US-10-619-906-5
Perfect score: 19
Sequence: 1 aattattgattctgtaggt 19

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	19	100.0	1410	3	US-08-068-392-1	Sequence 1, Appli
c	2	19	100.0	1410	3	US-08-396-988-1	Sequence 1, Appli
c	3	19	100.0	1780	4	US-09-949-016-5005	Sequence 5005, Ap
c	4	17	89.5	47	4	US-09-422-978-2352	Sequence 2352, Ap
c	5	16	84.2	597	4	US-09-248-796A-3451	Sequence 3451, Ap
c	6	16	84.2	1572	4	US-09-248-796A-3452	Sequence 3452, Ap
	7	15	78.9	498	4	US-09-854-133-345	Sequence 345, App
c	8	15	78.9	601	4	US-09-949-016-69808	Sequence 69808, A
	9	15	78.9	2070	4	US-09-949-016-2056	Sequence 2056, Ap
c	10	15	78.9	21125	4	US-09-949-016-15108	Sequence 15108, A
	11	15	78.9	39318	4	US-09-949-016-13798	Sequence 13798, A
c	12	15	78.9	51161	4	US-09-949-016-17416	Sequence 17416, A
	13	14	73.7	315	3	US-09-134-001C-250	Sequence 250, App
	14	14	73.7	373	3	US-09-615-192A-357	Sequence 357, App
	15	14	73.7	594	4	US-09-328-352-3288	Sequence 3288, Ap
c	16	14	73.7	601	4	US-09-949-016-27232	Sequence 27232, A
	17	14	73.7	601	4	US-09-949-016-81772	Sequence 81772, A
c	18	14	73.7	601	4	US-09-949-016-122846	Sequence 122846,
c	19	14	73.7	601	4	US-09-949-016-130732	Sequence 130732,
c	20	14	73.7	601	4	US-09-949-016-160756	Sequence 160756,
	21	14	73.7	900	4	US-09-543-681A-4127	Sequence 4127, Ap
c	22	14	73.7	987	4	US-09-248-796A-1645	Sequence 1645, Ap
	23	14	73.7	1131	4	US-09-248-796A-4531	Sequence 4531, Ap
c	24	14	73.7	1247	3	US-09-178-115-110	Sequence 110, App
c	25	14	73.7	1247	3	US-09-177-776-110	Sequence 110, App
c	26	14	73.7	1509	4	US-09-134-000C-2709	Sequence 2709, Ap
c	27	14	73.7	1617	2	US-08-467-963C-25	Sequence 25, Appl
c	28	14	73.7	1617	2	US-08-838-189D-25	Sequence 25, Appl
c	29	14	73.7	1617	3	US-08-852-344D-25	Sequence 25, Appl
c	30	14	73.7	1617	3	US-08-344-639E-25	Sequence 25, Appl
c	31	14	73.7	1844	2	US-08-467-963C-1	Sequence 1, Appli
c	32	14	73.7	1844	2	US-08-838-189D-1	Sequence 1, Appli
c	33	14	73.7	1844	3	US-08-852-344D-1	Sequence 1, Appli
c	34	14	73.7	1844	3	US-08-344-639E-1	Sequence 1, Appli
c	35	14	73.7	1844	3	US-08-467-969A-1	Sequence 1, Appli
c	36	14	73.7	1844	3	US-08-467-961A-1	Sequence 1, Appli
c	37	14	73.7	1844	3	US-08-001-554A-1	Sequence 1, Appli
	38	14	73.7	2007	4	US-09-617-594A-3	Sequence 3, Appli
	39	14	73.7	2354	4	US-09-023-655-1080	Sequence 1080, Ap
c	40	14	73.7	2501	3	US-08-787-739-58	Sequence 58, Appl
c	41	14	73.7	2501	3	US-09-178-115-58	Sequence 58, Appl
c	42	14	73.7	2501	3	US-09-177-776-58	Sequence 58, Appl
	43	14	73.7	2966	4	US-09-174-937-7	Sequence 7, Appli
c	44	14	73.7	3272	4	US-09-710-279-3700	Sequence 3700, Ap
c	45	14	73.7	3532	3	US-08-787-739-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1
US-08-068-392-1/c
; Sequence 1, Application US/08068392
; Patent No. 6150152
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,392
; FILING DATE: 19930528
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25275
; REFERENCE/DOCKET NUMBER: 07-24(12406)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1410 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1410
US-08-068-392-1

Query Match 100.0%; Score 19; DB 3; Length 1410;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 361 AATTATTGATTCTGTAGGT 343

```
RESULT 2
US-08-396-988-1/c
; Sequence 1, Application US/08396988
; Patent No. 6204043
;  GENERAL INFORMATION:
;    APPLICANT:  Shapiro, Steven M.
;    TITLE OF INVENTION:  Human Macrophage Metalloproteinase
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scott J. Meyer, Monsanto Co., A3SM
;      STREET:  800 N. Lindbergh Blvd.
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63167
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/396,988
;      FILING DATE:  01-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/068,392
;      FILING DATE:  28-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyer, Scott J.
;      REGISTRATION NUMBER:  25275
;      REFERENCE/DOCKET NUMBER:  07-24(12406)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)694-3117
;      TELEFAX:  (314)694-5435
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1410 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1410
US-08-396-988-1

  Query Match             100.0%;  Score 19;  DB 3;  Length 1410;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 3
US-09-949-016-5005/c
; Sequence 5005, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5005
;   LENGTH: 1780
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-5005

  Query Match             100.0%;  Score 19;  DB 4;  Length 1780;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 4
US-09-422-978-2352/c
; Sequence 2352, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2352
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-10570-107 : polymorphic base G or A
US-09-422-978-2352

  Query Match              89.5%;  Score 17;  DB 4;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db         21 TTATTGATTCTGTAGGT 5


RESULT 5
US-09-248-796A-3451/c
; Sequence 3451, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3451
```

; LENGTH: 597
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3451

  Query Match 84.2%; Score 16; DB 4; Length 597;
  Best Local Similarity 100.0%; Pred. No. 6.1;
  Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TATTGATTCTGTAGGT 19
      ||||||||||||||||
Db 39 TATTGATTCTGTAGGT 24

RESULT 6
US-09-248-796A-3452/c
; Sequence 3452, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3452
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3452

  Query Match 84.2%; Score 16; DB 4; Length 1572;
  Best Local Similarity 100.0%; Pred. No. 5.7;
  Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TATTGATTCTGTAGGT 19
      ||||||||||||||||
Db 45 TATTGATTCTGTAGGT 30

RESULT 7
US-09-854-133-345
; Sequence 345, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 345
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-345

  Query Match 78.9%; Score 15; DB 4; Length 498;
  Best Local Similarity 100.0%; Pred. No. 22;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGTA 16
      |||||||||||||||
Db 340 ATTATTGATTCTGTA 354

RESULT 8
US-09-949-016-69808/c
; Sequence 69808, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69808
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-69808

  Query Match 78.9%; Score 15; DB 4; Length 601;
  Best Local Similarity 100.0%; Pred. No. 21;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGTA 16
      |||||||||||||||
Db 381 ATTATTGATTCTGTA 367

RESULT 9
US-09-949-016-2056
; Sequence 2056, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2056
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2056

  Query Match 78.9%; Score 15; DB 4; Length 2070;
  Best Local Similarity 100.0%; Pred. No. 20;
  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGTA 16
      |||||||||||||||
Db 249 ATTATTGATTCTGTA 263

RESULT 10

US-09-949-016-15108/c
; Sequence 15108, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15108
; LENGTH: 21125
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(21125)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15108

Query Match 78.9%; Score 15; DB 4; Length 21125;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TATTGATTCTGTAGG 18
||||||||||||||||
Db 9015 TATTGATTCTGTAGG 9001

RESULT 11
US-09-949-016-13798
; Sequence 13798, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13798
; LENGTH: 39318
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13798

Query Match 78.9%; Score 15; DB 4; Length 39318;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGTA 16
||||||||||||||||
Db 2249 ATTATTGATTCTGTA 2263

RESULT 12
US-09-949-016-17416/c
; Sequence 17416, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17416
; LENGTH: 51161
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17416

Query Match 78.9%; Score 15; DB 4; Length 51161;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTAG 17
|||||||||||||||
Db 328 TTATTGATTCTGTAG 314

RESULT 13
US-09-134-001C-250
; Sequence 250, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 250
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-250

Query Match 73.7%; Score 14; DB 3; Length 315;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTA 16
||||||||||||||
Db 209 TTATTGATTCTGTA 222

RESULT 14
US-09-615-192A-357
; Sequence 357, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U

; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 373
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-615-192A-357

Query Match 73.7%; Score 14; DB 3; Length 373;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGT 15
||||||||||||||
Db 344 ATTATTGATTCTGT 357

RESULT 15
US-09-328-352-3288
; Sequence 3288, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3288
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3288

Query Match 73.7%; Score 14; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTA 16
||||||||||||||
Db 106 TTATTGATTCTGTA 119

RESULT 16
US-09-949-016-27232/c
; Sequence 27232, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27232
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27232

Query Match 73.7%; Score 14; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTG 14
||||||||||||||
Db 57 AATTATTGATTCTG 44

RESULT 17
US-09-949-016-81772
; Sequence 81772, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81772
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-81772

Query Match 73.7%; Score 14; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTG 14
||||||||||||||
Db 506 AATTATTGATTCTG 519

RESULT 18
US-09-949-016-122846/c
; Sequence 122846, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122846
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-122846

Query Match 73.7%; Score 14; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTG 14
              ||||||||||||||
Db        434 AATTATTGATTCTG 421
```

RESULT 19
US-09-949-016-130732/c
; Sequence 130732, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 130732
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-130732

Query Match 73.7%; Score 14; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTG 14
              ||||||||||||||
Db        434 AATTATTGATTCTG 421
```

RESULT 20
US-09-949-016-160756/c
; Sequence 160756, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 160756
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-160756

Query Match 73.7%; Score 14; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTG 14
              ||||||||||||||
```

```
Db         57 AATTATTGATTCTG 44
```

RESULT 21
US-09-543-681A-4127
; Sequence 4127, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4127
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-4127

Query Match 73.7%; Score 14; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTG 14
              ||||||||||||||
Db        717 AATTATTGATTCTG 730
```

RESULT 22
US-09-248-796A-1645/c
; Sequence 1645, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1645
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1645

Query Match 73.7%; Score 14; DB 4; Length 987;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 TATTGATTCTGTAG 17
              ||||||||||||||
Db        613 TATTGATTCTGTAG 600
```

RESULT 23
US-09-248-796A-4531
; Sequence 4531, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4531
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4531

Query Match 73.7%; Score 14; DB 4; Length 1131;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTG 14
||||||||||||||
Db 354 AATTATTGATTCTG 367

RESULT 24
US-09-178-115-110/c
; Sequence 110, Application US/09178115
; Patent No. 6297041
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/178,115
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 09/177,776
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 110
; LENGTH: 1247
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1247)
US-09-178-115-110

Query Match 73.7%; Score 14; DB 3; Length 1247;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGT 15
||||||||||||||
Db 810 ATTATTGATTCTGT 797

RESULT 25
US-09-177-776-110/c
; Sequence 110, Application US/09177776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/177,776A
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 110
; LENGTH: 1247
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1247)
US-09-177-776-110

Query Match 73.7%; Score 14; DB 3; Length 1247;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGT 15
||||||||||||||
Db 810 ATTATTGATTCTGT 797

RESULT 26
US-09-134-000C-2709/c
; Sequence 2709, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2709
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2709

Query Match 73.7%; Score 14; DB 4; Length 1509;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 TTGATTCTGTAGGT 19
||||||||||||||
Db 46 TTGATTCTGTAGGT 33

RESULT 27
US-08-467-963C-25/c
; Sequence 25, Application US/08467963C
; Patent No. 5968776
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,963C
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/838,189
; FILING DATE: 16-APR-1997
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1617 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-467-963C-25

Query Match 73.7%; Score 14; DB 2; Length 1617;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTA 16
||||||||||||||
Db 787 TTATTGATTCTGTA 774

RESULT 28
US-08-838-189D-25/c
; Sequence 25, Application US/08838189D
; Patent No. 5998169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,189D
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1617 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-838-189D-25

Query Match 73.7%; Score 14; DB 2; Length 1617;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTA 16
||||||||||||||
Db 787 TTATTGATTCTGTA 774

RESULT 29
US-08-852-344D-25/c
; Sequence 25, Application US/08852344D
; Patent No. 6017539
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION

; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
; TITLE OF INVENTION: SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,344D
; FILING DATE: 07-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,639
; FILING DATE: 14-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1617 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-852-344D-25

Query Match 73.7%; Score 14; DB 3; Length 1617;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTA 16
||||||||||||||
Db 787 TTATTGATTCTGTA 774

RESULT 30
US-08-344-639E-25/c
; Sequence 25, Application US/08344639E
; Patent No. 6033668
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasyshyn, Mary E
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
; TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,639E
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-391 MIS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1617 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-344-639E-25

Query Match 73.7%; Score 14; DB 3; Length 1617;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTA 16
||||||||||||||
Db 787 TTATTGATTCTGTA 774

RESULT 31
US-08-467-963C-1/c
; Sequence 1, Application US/08467963C
; Patent No. 5968776
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EWASYSHYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,963C
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/838,189
; FILING DATE: 16-APR-1997
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEWART, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-474 MIS:jb
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1844 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-467-963C-1

  Query Match             73.7%;  Score 14;  DB 2;  Length 1844;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db        970 TTATTGATTCTGTA 957


RESULT 32
US-08-838-189D-1/c
; Sequence 1, Application US/08838189D
; Patent No. 5998169
;  GENERAL INFORMATION:
;    APPLICANT:  KLEIN, Michel H
;    APPLICANT:  DU, Run-Pan
;    APPLICANT:  EWASYSHYN, Mary E
;    TITLE OF INVENTION:  MULTIMERIC HYBRID GENE ENCODING A
;    TITLE OF INVENTION:  CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
;    TITLE OF INVENTION:  PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  6th Floor, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/838,189D
;      FILING DATE:  16-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/001,554
;      FILING DATE:  06-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9200117.1
;      FILING DATE:  06-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEWART, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-687 MIS:jb
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1844 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-838-189D-1

  Query Match             73.7%;  Score 14;  DB 2;  Length 1844;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db        970 TTATTGATTCTGTA 957


RESULT 33
US-08-852-344D-1/c
; Sequence 1, Application US/08852344D
; Patent No. 6017539
;  GENERAL INFORMATION:
;    APPLICANT:  KLEIN, Michel H
;    APPLICANT:  DU, Run-Pan
;    APPLICANT:  EWASYSHYN, Mary E
;    TITLE OF INVENTION:  CHIMERIC PROTEIN WHICH CONFERS PROTECTION
;    TITLE OF INVENTION:  AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
;    TITLE OF INVENTION:  SYNCYTIAL VIRUS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  6th Floor, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,344D
;      FILING DATE:  07-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/344,639
;      FILING DATE:  14-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9200117.1
;      FILING DATE:  06-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEWART, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-688 MIS:jb
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1844 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-852-344D-1

  Query Match             73.7%;  Score 14;  DB 3;  Length 1844;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db        970 TTATTGATTCTGTA 957


RESULT 34
US-08-344-639E-1/c
; Sequence 1, Application US/08344639E
```

```
; Patent No. 6033668
;  GENERAL INFORMATION:
;    APPLICANT:  Klein, Michel H
;    APPLICANT:  Du, Run-Pan
;    APPLICANT:  Ewasyshyn, Mary E
;    TITLE OF INVENTION:  CHIMERIC PROTEIN WHICH CONFERS
;    TITLE OF INVENTION:  PROTECTION AGAINST PARAINFLUENZA VIRUS
;    TITLE OF INVENTION:  AND RESPIRATORY SYNCYTIAL VIRUS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  330 University Avenue, 6th Floor
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/344,639E
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/001,554
;      FILING DATE:  06-JAN-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9200117.1
;      FILING DATE:  06-JAN-1992
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-391 MIS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;      TELEX:  065-24567 SIMBAS
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1844 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-344-639E-1

  Query Match             73.7%;  Score 14;  DB 3;  Length 1844;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db        970 TTATTGATTCTGTA 957


RESULT 35
US-08-467-969A-1/c
; Sequence 1, Application US/08467969A
; Patent No. 6168786
;  GENERAL INFORMATION:
;    APPLICANT:  Klein, Michel H
;    APPLICANT:  Du, Run-Pan
;    APPLICANT:  Ewasyshyn, Mary E
;    TITLE OF INVENTION:  Chimeric Immunogens
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  330 University Avenue, 6th Floor
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,969A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/001,554
;      FILING DATE:  06-JAN-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9200117.1
;      FILING DATE:  06-JAN-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-475 MIS:bh
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;      TELEX:  065-24567 SIMBAS
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1844 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-467-969A-1

  Query Match             73.7%;  Score 14;  DB 3;  Length 1844;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db        970 TTATTGATTCTGTA 957


RESULT 36
US-08-467-961A-1/c
; Sequence 1, Application US/08467961A
; Patent No. 6171783
;  GENERAL INFORMATION:
;    APPLICANT:  Klein, Michel H
;    APPLICANT:  Du, Run-Pan
;    APPLICANT:  Ewasyshyn, Mary E
;    TITLE OF INVENTION:  Chimeric Immunogens
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  330 University Avenue, 6TH Floor
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,961A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/001,554
;      FILING DATE:  06-JAN-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9200117.1
;      FILING DATE:  06-JAN-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-476 MIS:bh
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;      TELEX:  065-24567 SIMBAS
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1844 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-467-961A-1

  Query Match             73.7%;  Score 14;  DB 3;  Length 1844;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db        970 TTATTGATTCTGTA 957


RESULT 37
US-08-001-554A-1/c
; Sequence 1, Application US/08001554A
; Patent No. 6225091
;  GENERAL INFORMATION:
;    APPLICANT:  Klein, Michel H
;    APPLICANT:  Du, Run-Pan
;    APPLICANT:  Ewasyshyn, Mary E
;    TITLE OF INVENTION:  Chimeric Immunogens
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  330 University Avenue, 6th Floor
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/001,554A
;      FILING DATE:  06-JAN-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-286
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;      TELEX:  065-24567 SIMBAS
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1844 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-001-554A-1

  Query Match             73.7%;  Score 14;  DB 3;  Length 1844;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db        970 TTATTGATTCTGTA 957


RESULT 38
US-09-617-594A-3
; Sequence 3, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
;  APPLICANT: Audonnet, et al.
;  TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT VI
;  FILE REFERENCE: 454313-3151.1
;  CURRENT APPLICATION NUMBER: US/09/617,594A
;  CURRENT FILING DATE:  2000-07-14
;  PRIOR APPLICATION NUMBER: 60/193,332
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: France 00 01761
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: France 99 09421
;  PRIOR FILING DATE: 1999-07-16
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 2007
;   TYPE: DNA
;   ORGANISM: Feline calcivirus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2004)
;   OTHER INFORMATION: coding sequence of the "capsid" protein of FCV 431 strain
US-09-617-594A-3

  Query Match             73.7%;  Score 14;  DB 4;  Length 2007;
  Best Local Similarity   100.0%;  Pred. No. 69;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTG 14
              ||||||||||||||
Db       1830 AATTATTGATTCTG 1843


RESULT 39
US-09-023-655-1080
; Sequence 1080, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/09/023,655
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0001 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1080:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2354 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g182573
US-09-023-655-1080

  Query Match             73.7%;  Score 14;  DB 4;  Length 2354;
  Best Local Similarity   100.0%;  Pred. No. 69;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTA 16
              ||||||||||||||
Db       2311 TTATTGATTCTGTA 2324


RESULT 40
US-08-787-739-58/c
; Sequence 58, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street, Suite 610
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/787,739
;      FILING DATE:  24-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,049
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/486,756
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,658
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,862
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,863
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,077
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-981-2034
;      TELEFAX:  415-981-0332
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2501 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;      DESCRIPTION:  MN genomic region between pMN1 and
;      DESCRIPTION:  Bd3 assumed to contain regulatory
;      DESCRIPTION:  elements
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-787-739-58

  Query Match             73.7%;  Score 14;  DB 3;  Length 2501;
  Best Local Similarity   100.0%;  Pred. No. 68;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ATTATTGATTCTGT 15
              ||||||||||||||
Db       2131 ATTATTGATTCTGT 2118


RESULT 41
US-09-178-115-58/c
; Sequence 58, Application US/09178115
; Patent No. 6297041
; GENERAL INFORMATION:
;  APPLICANT: Zavada, Jan
;  APPLICANT: Pastorekova, Silvia
;  APPLICANT: Pastorek, Jaromir
;  TITLE OF INVENTION: MN Gene and Protein
;  FILE REFERENCE: D-0021.5A
;  CURRENT APPLICATION NUMBER: US/09/178,115
;  CURRENT FILING DATE: 1998-10-23
;  EARLIER APPLICATION NUMBER: 09/177,776
;  EARLIER FILING DATE: 1998-10-23
;  EARLIER APPLICATION NUMBER: 08/787,739
;  EARLIER FILING DATE: 1997-01-24
;  EARLIER APPLICATION NUMBER: 08/485,049
;  EARLIER FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 08/486,756
;  EARLIER FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 08/477,504
;  EARLIER FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 08/481,658
;  EARLIER FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 08/485,862
;  EARLIER FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 08/485,863
;  EARLIER FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 08/487,077
;  EARLIER FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 08/260,190
;  EARLIER FILING DATE: 1994-06-15
;  EARLIER APPLICATION NUMBER: 08/177,093
;  EARLIER FILING DATE: 1993-12-30
;  EARLIER APPLICATION NUMBER: 07/964,589
;  EARLIER FILING DATE: 1992-10-21
```

; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2501)
US-09-178-115-58

Query Match 73.7%; Score 14; DB 3; Length 2501;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGT 15
||||||||||||||
Db 2131 ATTATTGATTCTGT 2118

RESULT 42
US-09-177-776-58/c
; Sequence 58, Application US/09177776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/177,776A
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2501)
US-09-177-776-58

Query Match 73.7%; Score 14; DB 3; Length 2501;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGT 15
||||||||||||||
Db 2131 ATTATTGATTCTGT 2118

RESULT 43
US-09-174-937-7
; Sequence 7, Application US/09174937
; Patent No. 6756212
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: Isolated Proteins and Nucleic Acid Molecules Having
; TITLE OF INVENTION: Homology to NIP2 Proteins and Uses Thereof
; FILE REFERENCE: mni061
; CURRENT APPLICATION NUMBER: US/09/174,937
; CURRENT FILING DATE: 1998-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(433)
US-09-174-937-7

Query Match 73.7%; Score 14; DB 4; Length 2966;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 TTGATTCTGTAGGT 19
||||||||||||||
Db 1787 TTGATTCTGTAGGT 1800

RESULT 44
US-09-710-279-3700/c
; Sequence 3700, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3700
; LENGTH: 3272
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3700

Query Match 73.7%; Score 14; DB 4; Length 3272;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTA 16
||||||||||||||
Db 1769 TTATTGATTCTGTA 1756

RESULT 45
US-08-787-739-90/c
; Sequence 90, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; DESCRIPTION: position 7 to position 3538 of
; DESCRIPTION: Figure 2A-2F
US-08-787-739-90

Query Match 73.7%; Score 14; DB 3; Length 3532;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 ATTATTGATTCTGT 15
              ||||||||||||||
Db       2131 ATTATTGATTCTGT 2118
```

Search completed: April 26, 2005, 16:56:41
Job time : 139 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 26, 2005, 11:52:57 ; Search time 1816 Seconds
(without alignments)
506.965 Million cell updates/sec

Title: US-10-619-906-5
Perfect score: 19
Sequence: 1 aattattgattctgtaggt 19

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	19	100.0	458	6	AX393790	AX393790 Sequence
c	2	19	100.0	1410	6	AR143096	AR143096 Sequence
c	3	19	100.0	1653	6	AX481487	AX481487 Sequence
c	4	19	100.0	1778	6	CQ833974	CQ833974 Sequence
c	5	19	100.0	1778	6	AX329930	AX329930 Sequence
c	6	19	100.0	1778	6	AX330492	AX330492 Sequence
c	7	19	100.0	1778	6	AX332571	AX332571 Sequence
c	8	19	100.0	1778	6	AX333645	AX333645 Sequence
c	9	19	100.0	1778	6	AX334317	AX334317 Sequence
c	10	19	100.0	1778	6	AX393829	AX393829 Sequence
c	11	19	100.0	1778	6	AX409532	AX409532 Sequence
c	12	19	100.0	1778	6	AX663647	AX663647 Sequence
c	13	19	100.0	1778	9	HUMHME	L23808 Human metal
c	14	19	100.0	1780	6	CQ730208	CQ730208 Sequence
c	15	17	89.5	47	6	AR290617	AR290617 Sequence
	16	17	89.5	1146	1	BCE304815	AJ304815 Bacillus
	17	17	89.5	1146	1	BCE304816	AJ304816 Bacillus
c	18	17	89.5	73353	2	AC101422	AC101422 Mus muscu
c	19	17	89.5	87206	9	HS188H10	AL021451 Human DNA
c	20	17	89.5	92611	9	AC006398	AC006398 Homo sapi
c	21	17	89.5	101443	9	AC091621	AC091621 Papio anu
c	22	17	89.5	110000	2	LMFLCHR31_02	Continuation (3 of
	23	17	89.5	140805	2	AC092456	AC092456 Homo sapi
	24	17	89.5	150070	2	AC007670	AC007670 Homo sapi
c	25	17	89.5	153270	2	AC091661	AC091661 Pan trogl
c	26	17	89.5	154436	9	AC098817	AC098817 Homo sapi
	27	17	89.5	157800	2	AL357074	AL357074 Homo sapi
c	28	17	89.5	159790	2	AC013613	AC013613 Homo sapi
c	29	17	89.5	164958	9	AC099340	AC099340 Homo sapi
c	30	17	89.5	172263	2	AC119280	AC119280 Mus muscu
c	31	17	89.5	182644	9	AC098811	AC098811 Papio anu
	32	17	89.5	183774	2	AC102150	AC102150 Mus muscu
	33	17	89.5	194648	2	AC133083	AC133083 Mus muscu
c	34	17	89.5	196915	10	AL929408	AL929408 Mouse DNA
	35	17	89.5	199373	2	AC119027	AC119027 Rattus no
	36	17	89.5	204308	2	AC138676	AC138676 Mus muscu
c	37	17	89.5	210358	2	AC148095	AC148095 Mus muscu
c	38	17	89.5	213348	10	AC134898	AC134898 Mus muscu
c	39	17	89.5	221618	2	AC098411	AC098411 Rattus no
	40	17	89.5	235315	2	AC103654	AC103654 Mus muscu
c	41	17	89.5	241151	2	AC125675	AC125675 Rattus no
	42	17	89.5	256265	2	AC122943	AC122943 Rattus no
	43	17	89.5	273331	2	AC110676	AC110676 Rattus no
	44	17	89.5	292913	2	AC106456	AC106456 Rattus no
c	45	17	89.5	300854	1	AE017014	AE017014 Bacillus

ALIGNMENTS

RESULT 1
AX393790/c
LOCUS AX393790 458 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 179 from Patent WO0212329.
ACCESSION AX393790
VERSION AX393790.1 GI:19701757
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Wang,T. and Fan,L.
TITLE Compositions and methods for the therapy and diagnosis of head and neck cancer
JOURNAL Patent: WO 0212329-A 179 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .458
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 458;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        348 AATTATTGATTCTGTAGGT 330
```

RESULT 2
AR143096/c
LOCUS AR143096 1410 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6204043.
ACCESSION AR143096
VERSION AR143096.1 GI:15104382
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.

```
            Unclassified.
REFERENCE   1  (bases 1 to 1410)
  AUTHORS   Shapiro,S.D.
  TITLE     Human macrophage metalloproteinase
  JOURNAL   Patent: US 6204043-A 1 20-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .1410
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1410;
  Best Local Similarity   100.0%;  Pred. No. 5.2;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 3
AX481487/c
LOCUS       AX481487               1653 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 101 from Patent WO02055693.
ACCESSION   AX481487
VERSION     AX481487.1  GI:22316401
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kreutzer,R., Limmer,S., Rost,S. and Hadwiger,P.
  TITLE     Method for inhibiting the expression of a target gene
  JOURNAL   Patent: WO 02055693-A 101 18-JUL-2002;
            Ribopharma AG (DE)
FEATURES             Location/Qualifiers
     source          1. .1653
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1653;
  Best Local Similarity   100.0%;  Pred. No. 5.1;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 4
CQ833974/c
LOCUS       CQ833974               1778 bp    DNA     linear   PAT 29-JUL-2004
DEFINITION  Sequence 10 from Patent EP1439393.
ACCESSION   CQ833974
VERSION     CQ833974.1  GI:50833579
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Astle,J.H., Boardman,L.A., Bugart,L.J., Burgess,C.C., Catino,T.J.,
            Dwivedi,P., Huntress,M., Johnson,K.A., Lewis,M.E., Maimonis,P.J.,
            Myerow,S.H., Brown-Shimer,S.L., Thiagalingam,A., Thibodeau,S.N. and
            Molino,G.A.
  TITLE     Detection methods using TIMP 1 for colon cancer diagnosis
  JOURNAL   Patent: EP 1439393-A 10 21-JUL-2004;
            Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION
            AND RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 5
AX329930/c
LOCUS       AX329930               1778 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 439 from Patent WO0194629.
ACCESSION   AX329930
VERSION     AX329930.1  GI:18102908
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 439 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 6
AX330492/c
LOCUS       AX330492               1778 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 1001 from Patent WO0194629.
ACCESSION   AX330492
VERSION     AX330492.1  GI:18103470
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 1001 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 7
AX332571/c
LOCUS       AX332571                1778 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 3080 from Patent WO0194629.
ACCESSION   AX332571
VERSION     AX332571.1  GI:18123205
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 3080 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 8
AX333645/c
LOCUS       AX333645                1778 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 4154 from Patent WO0194629.
ACCESSION   AX333645
VERSION     AX333645.1  GI:18124364
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 4154 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 9
AX334317/c
LOCUS       AX334317                1778 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 4826 from Patent WO0194629.
ACCESSION   AX334317
VERSION     AX334317.1  GI:18125036
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 4826 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 10
AX393829/c
LOCUS       AX393829                1778 bp    DNA     linear   PAT 23-MAR-2002
DEFINITION  Sequence 218 from Patent WO0212329.
ACCESSION   AX393829
VERSION     AX393829.1  GI:19701796
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T. and Fan,L.
  TITLE     Compositions and methods for the therapy and diagnosis of head and
            neck cancer
  JOURNAL   Patent: WO 0212329-A 218 14-FEB-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 11
```

```
AX409532/c
LOCUS       AX409532                1778 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 2179 from Patent WO0229103.
ACCESSION   AX409532
VERSION     AX409532.1  GI:21442237
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2179 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. L23808"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 12
AX663647/c
LOCUS       AX663647                1778 bp    DNA     linear   PAT 22-MAR-2003
DEFINITION  Sequence 22 from Patent WO02097127.
ACCESSION   AX663647
VERSION     AX663647.1  GI:29163855
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Oellers,N., Gehrmann,M., Kallabis,H., Hall,R., Schulze,T. and
            Kroegel,C.
  TITLE     Genes and proteins for prevention, prediction, diagnosis, prognosis
            and treatment of chronic lung disease
  JOURNAL   Patent: WO 02097127-A 22 05-DEC-2002;
            Bayer Aktiengesellschaft (DE)
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 13
HUMHME/c
LOCUS       HUMHME                  1778 bp    mRNA    linear   PRI 31-DEC-1994
DEFINITION  Human metalloproteinase (HME) mRNA, complete cds.
ACCESSION   L23808
VERSION     L23808.1  GI:435969
KEYWORDS    metalloproteinase.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1778)
  AUTHORS   Shapiro,S.D., Kobayashi,D.K. and Ley,T.J.
  TITLE     Cloning and characterization of a unique elastolytic
            metalloproteinase produced by human alveolar macrophages
  JOURNAL   J. Biol. Chem. 268 (32), 23824-23829 (1993)
  MEDLINE   94043200
   PUBMED   8226919
COMMENT     Original source text: Homo sapiens adult cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .1778
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_type="macrophage"
                     /dev_stage="adult"
     gene            1. .1778
                     /gene="HME"
     CDS             13. .1425
                     /gene="HME"
                     /codon_start=1
                     /product="metalloproteinase"
                     /protein_id="AAA58658.1"
                     /db_xref="GI:435970"
                     /translation="MKFLLILLLQATASGALPLNSSTSLEKNNVLFGERYLEKFYGLE
                     INKLPVTKMKYSGNLMKEKIQEMQHFLGLKVTGQLDTSTLEMMHAPRCGVPDLHHFRE
                     MPGGPVWRKHYITYRINNYTPDMNREDVDYAIRKAFQVWSNVTPLKFSKINTGMADIL
                     VVFARGAHGDFHAFDGKGGILAHAFGPGSGIGGDAHFDEDEFWTTHSGGTNLFLTAVH
                     EIGHSLGLGHSSDPKAVMFPTYKYVDINTFRLSADDIRGIQSLYGDPKENQRLPNPDN
                     SEPALCDPNLSFDAVTTVGNKIFFFKDRFFWLKVSERPKTSVNLISSLWPTLPSGIEA
                     AYEIEARNQVFLFKDDKYWLISNLRPEPNYPKSIHSFGFPNFVKKIDAAVFNPRFYRT
                     YFFVDNQYWRYDERRQMMDPGYPKLITKNFQGIGPKIDAVFYSKNKYYYFFQGSNQFE
                     YDFLLQRITKTLKSNSWFGC"
     polyA_site      1778
                     /gene="HME"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 14
CQ730208/c
LOCUS       CQ730208                1780 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 16142 from Patent WO02068579.
ACCESSION   CQ730208
VERSION     CQ730208.1  GI:42303435
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 16142 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1780
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             100.0%;  Score 19;  DB 6;  Length 1780;
  Best Local Similarity   100.0%;  Pred. No. 5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 15
AR290617/c
LOCUS       AR290617                 47 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 2352 from patent US 6537751.
ACCESSION   AR290617
VERSION     AR290617.1  GI:31677901
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 2352 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 6;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db         21 TTATTGATTCTGTAGGT 5


RESULT 16
BCE304815
LOCUS       BCE304815               1146 bp    DNA     linear   BCT 01-JUL-2001
DEFINITION  Bacillus cereus partial clo1 gene for putative cereolysin O.
ACCESSION   AJ304815
VERSION     AJ304815.1  GI:14589120
KEYWORDS    cereolysin O; clo1 gene.
SOURCE      Bacillus cereus ATCC 14579
  ORGANISM  Bacillus cereus ATCC 14579
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Michelet,N.
  TITLE     Ubiquity and multicopy of the cereolysin O gene among members of
            the Bacillus cereus sensu lato group
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1146)
  AUTHORS   Mahillon,J.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-2000) Mahillon J.G., Laboratory of Microbial
            Genetics, Catholic University of Louvain, Place Croix du Sud, 2/12,
            B-1348, BELGIUM
FEATURES             Location/Qualifiers
     source          1. .1146
                     /organism="Bacillus cereus ATCC 14579"
                     /mol_type="genomic DNA"
                     /strain="ATCC14579"
                     /isolate="Type strain"
                     /db_xref="taxon:226900"
                     /clone="ATCC.1"
                     /country="USA"
     gene            1. .1146
                     /gene="clo1"
     CDS             <1. .>1146
                     /gene="clo1"
                     /function="hemolysis"
                     /codon_start=1
                     /transl_table=11
                     /product="putative cereolysin O"
                     /protein_id="CAC43083.1"
                     /db_xref="GI:14589121"
                     /db_xref="GOA:Q93LB0"
                     /db_xref="UniProt/TrEMBL:Q93LB0"
                     /translation="TSPVDISIIDSVVNRTYPGAVQLANKAFADNQPSLLVAKRKPLN
                     ISIDLPGMRKENTITVQNPTYGNVAGAVDDLVSTWNEKYSTTHTLPARMQYTESMVYS
                     KSQIASALNVNAKYLDNSLNIDFNAVANGEKKVMVAAYKQIFYTVSAELPNNPSDLFG
                     NSVTFDELTRKGVSNSAPPVMVSNVAYGRTIYVKLETTSKSKDVQAAFKALLKDNSVE
                     TSGQYKDIFEESTFTAVVLGGDAKEHNKVVTKDFNEIRNIIKDNAELSLKNPAYPISY
                     TSTFLKDNATAAVHNNTDYIETTTTEYSSAKMTLDHYGAYVAQFDVSWDEFTFDQKGN
                     EVLTHKTWDGSGKDKTAHYSTVIPLPPNSKNIKIVARECTGLAWEWWR"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 1;  Length 1146;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db         21 AATTATTGATTCTGTAG 37


RESULT 17
BCE304816
LOCUS       BCE304816               1146 bp    DNA     linear   BCT 01-JUL-2001
DEFINITION  Bacillus cereus partial clo2 gene for putative cereolysin O.
ACCESSION   AJ304816
VERSION     AJ304816.1  GI:14589122
KEYWORDS    cereolysin O; clo2 gene.
SOURCE      Bacillus cereus ATCC 14579
  ORGANISM  Bacillus cereus ATCC 14579
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1
  AUTHORS   Michelet,N.
  TITLE     Ubiquity and multicopy of the cereolysin O gene among members of
            the Bacillus cereus sensu lato group
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1146)
  AUTHORS   Mahillon,J.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-2000) Mahillon J.G., Laboratory of Microbial
            Genetics, Catholic University of Louvain, Place Croix du Sud, 2/12,
            B-1348, BELGIUM
FEATURES             Location/Qualifiers
     source          1. .1146
                     /organism="Bacillus cereus ATCC 14579"
                     /mol_type="genomic DNA"
                     /strain="ATCC14579"
                     /isolate="Type strain"
                     /db_xref="taxon:226900"
                     /clone="ATCC.2"
                     /country="USA"
     gene            1. .1146
                     /gene="clo2"
     CDS             <1. .>1146
                     /gene="clo2"
                     /function="hemolysis"
                     /codon_start=1
                     /transl_table=11
                     /product="putative cereolysin O"
                     /protein_id="CAC43084.1"
                     /db_xref="GI:14589123"
                     /db_xref="GOA:Q93LA9"
                     /db_xref="UniProt/TrEMBL:Q93LA9"
                     /translation="TSPVDISIIDSVVNRTYPGAVQLANKAFADNQPSLLVAKRKPLN
                     ISIDLPGMRKENTITVQNPTYGNVAGAVDDLVSTWNEKYSTTHTLPARMQYTESMVYS
```

KSQIASALNVYAKYLDNSLNIDFNAVANGEKKAMVAAYKQIFYTVSAELPNNPSDLFD
NSVTFDELTRKGVSNSAPPVMVSNVAYGRTIYVKLETTSKSKDVQAAFKALLKDNSVE
TSGQYKDIFEESTFTAVVLGGDAKEHNKVVTKDFNEIRNIIKDNAELSLKNPAYPISY
TSTFLKDNATAAVHNNTDYIETTTTEYSSAKMTLDHYGAYVAQFDVSWDEFTFDQKGN
EVLTHKTWDGSGKDKTAHYSTVIPLPPNSKNIKIVARECTGLAWEWWR"

ORIGIN

Query Match 89.5%; Score 17; DB 1; Length 1146;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAG 17
||||||||||||||||
Db 21 AATTATTGATTCTGTAG 37

RESULT 18
AC101422/c
LOCUS AC101422 73353 bp DNA linear HTG 23-NOV-2001
DEFINITION Mus musculus clone RP23-142G11, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC101422
VERSION AC101422.1 GI:17060197
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 73353)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-142G11
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 73353)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
Center project name: L16664
Center clone name: 142_G_11
--------------.
* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 676: contig of 676 bp in length
* 677 776: gap of 100 bp
* 777 1479: contig of 703 bp in length
* 1480 1579: gap of 100 bp
* 1580 2276: contig of 697 bp in length
* 2277 2376: gap of 100 bp
* 2377 3069: contig of 693 bp in length
* 3070 3169: gap of 100 bp
* 3170 3850: contig of 681 bp in length
* 3851 3950: gap of 100 bp
* 3951 4662: contig of 712 bp in length
* 4663 4762: gap of 100 bp
* 4763 5474: contig of 712 bp in length
* 5475 5574: gap of 100 bp
* 5575 6285: contig of 711 bp in length
* 6286 6385: gap of 100 bp
* 6386 7117: contig of 732 bp in length
* 7118 7217: gap of 100 bp
* 7218 7908: contig of 691 bp in length
* 7909 8008: gap of 100 bp
* 8009 8706: contig of 698 bp in length
* 8707 8806: gap of 100 bp
* 8807 9506: contig of 700 bp in length
* 9507 9606: gap of 100 bp
* 9607 10293: contig of 687 bp in length
* 10294 10393: gap of 100 bp
* 10394 11090: contig of 697 bp in length
* 11091 11190: gap of 100 bp
* 11191 11878: contig of 688 bp in length
* 11879 11978: gap of 100 bp
* 11979 12673: contig of 695 bp in length
* 12674 12773: gap of 100 bp
* 12774 13467: contig of 694 bp in length
* 13468 13567: gap of 100 bp
* 13568 14281: contig of 714 bp in length
* 14282 14381: gap of 100 bp
* 14382 15091: contig of 710 bp in length
* 15092 15191: gap of 100 bp
* 15192 15903: contig of 712 bp in length
* 15904 16003: gap of 100 bp
* 16004 16715: contig of 712 bp in length
* 16716 16815: gap of 100 bp
* 16816 17499: contig of 684 bp in length
* 17500 17599: gap of 100 bp
* 17600 18295: contig of 696 bp in length
* 18296 18395: gap of 100 bp
* 18396 19113: contig of 718 bp in length
* 19114 19213: gap of 100 bp
* 19214 19882: contig of 669 bp in length
* 19883 19982: gap of 100 bp
* 19983 20670: contig of 688 bp in length
* 20671 20770: gap of 100 bp
* 20771 21453: contig of 683 bp in length
* 21454 21553: gap of 100 bp
* 21554 22249: contig of 696 bp in length
* 22250 22349: gap of 100 bp
* 22350 23055: contig of 706 bp in length
* 23056 23155: gap of 100 bp
* 23156 23866: contig of 711 bp in length
* 23867 23966: gap of 100 bp
* 23967 24685: contig of 719 bp in length
* 24686 24785: gap of 100 bp
* 24786 25505: contig of 720 bp in length
* 25506 25605: gap of 100 bp
* 25606 26306: contig of 701 bp in length
* 26307 26406: gap of 100 bp

```
*    26407    27095: contig of 689 bp in length
*    27096    27195: gap of 100 bp
*    27196    27891: contig of 696 bp in length
*    27892    27991: gap of 100 bp
*    27992    28687: contig of 696 bp in length
*    28688    28787: gap of 100 bp
*    28788    29485: contig of 698 bp in length
*    29486    29585: gap of 100 bp
*    29586    30257: contig of 672 bp in length
*    30258    30357: gap of 100 bp
*    30358    31053: contig of 696 bp in length
*    31054    31153: gap of 100 bp
*    31154    31845: contig of 692 bp in length
*    31846    31945: gap of 100 bp
*    31946    32649: contig of 704 bp in length
*    32650    32749: gap of 100 bp
*    32750    33458: contig of 709 bp in length
*    33459    33558: gap of 100 bp
*    33559    34263: contig of 705 bp in length
*    34264    34363: gap of 100 bp
*    34364    35059: contig of 696 bp in length
*    35060    35159: gap of 100 bp
*    35160    35836: contig of 677 bp in length
*    35837    35936: gap of 100 bp
*    35937    36636: contig of 700 bp in length
*    36637    36736: gap of 100 bp
*    36737    37440: contig of 704 bp in length
*    37441    37540: gap of 100 bp
*    37541    38247: contig of 707 bp in length
*    38248    38347: gap of 100 bp
*    38348    39028: contig of 681 bp in length
*    39029    39128: gap of 100 bp
*    39129    39821: contig of 693 bp in length
*    39822    39921: gap of 100 bp
*    39922    40629: contig of 708 bp in length
*    40630    40729: gap of 100 bp
*    40730    41413: contig of 684 bp in length
*    41414    41513: gap of 100 bp
*    41514    42226: contig of 713 bp in length
*    42227    42326: gap of 100 bp
*    42327    43051: contig of 725 bp in length
*    43052    43151: gap of 100 bp
*    43152    43859: contig of 708 bp in length
*    43860    43959: gap of 100 bp
*    43960    44656: contig of 697 bp in length
*    44657    44756: gap of 100 bp
*    44757    45460: contig of 704 bp in length
*    45461    45560: gap of 100 bp
*    45561    46271: contig of 711 bp in length
*    46272    46371: gap of 100 bp
*    46372    47055: contig of 684 bp in length
*    47056    47155: gap of 100 bp
*    47156    47825: contig of 670 bp in length
*    47826    47925: gap of 100 bp
*    47926    48639: contig of 714 bp in length
*    48640    48739: gap of 100 bp
*    48740    49441: contig of 702 bp in length
*    49442    49541: gap of 100 bp
*    49542    50259: contig of 718 bp in length
*    50260    50359: gap of 100 bp
*    50360    51058: contig of 699 bp in length
*    51059    51158: gap of 100 bp
*    51159    51857: contig of 699 bp in length
*    51858    51957: gap of 100 bp
*    51958    52649: contig of 692 bp in length
*    52650    52749: gap of 100 bp
*    52750    53454: contig of 705 bp in length
*    53455    53554: gap of 100 bp
*    53555    54251: contig of 697 bp in length
*    54252    54351: gap of 100 bp
*    54352    55044: contig of 693 bp in length
*    55045    55144: gap of 100 bp
```

```
  Query Match             89.5%;  Score 17;  DB 2;  Length 73353;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      59266 TTATTGATTCTGTAGGT 59250


RESULT 19
HS188H10/c
LOCUS       HS188H10               87206 bp    DNA     linear   PRI 04-MAR-2003
DEFINITION  Human DNA sequence from clone RP1-188H10 on chromosome 6q22,
            complete sequence.
ACCESSION   AL021451
VERSION     AL021451.1  GI:2815069
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 87206)
  AUTHORS   Mashreghi-Mohammadi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     ---------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ----------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-188H10 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1. .87206
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704H10188"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="q22"
                     /clone="RP1-188H10"
                     /clone_lib="RPCI-1"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 9;  Length 87206;
```

```
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      30151 TTATTGATTCTGTAGGT 30135


RESULT 20
AC006398/c
LOCUS       AC006398               92611 bp    DNA     linear   PRI 08-OCT-2003
DEFINITION  Homo sapiens BAC clone CTA-348C20 from 7, complete sequence.
ACCESSION   AC006398
VERSION     AC006398.2  GI:21322241
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 92611)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 92611)
  AUTHORS   Bauer,C., Gibson,A. and Sun,H.
  TITLE     The sequence of Homo sapiens BAC clone CTA-348C20
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 92611)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 92611)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 92611)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 92611)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   7  (bases 1 to 92611)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   8  (bases 1 to 92611)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Jun 4, 2002 this sequence version replaced gi:4156132.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_RG348C20
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center. For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu

            SOURCE INFORMATION:
            Clone CTA-348C20 is from a release of the human BAC library
            CITB-HS-A. The library contains cloned DNA from human sperm. See:
            Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J.
            Kim et al., Genomics 34:213-8 (1996). The clone is available from
            Research Genetics, Inc. (http://www.resgen.com).
            VECTOR:  pBeloBAC11
            Selection:  chloramphenicol

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is CTA-471E18 the clone sequenced
            to the right is RP11-463M14.  Actual start of this clone is at base
            position 1 of CTA-348C20 actual end is at base position 92611 of
            CTA-348C20.
FEATURES            Location/Qualifiers
     source         1. .92611
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="7"
                    /map="7"
                    /clone="CTA-348C20"
                    /clone_lib="CITB-HS-A"
     repeat_region  300. .480
                    /rpt_family="MER1_type"
     repeat_region  536. .1057
                    /rpt_family="MER1_type"
     repeat_region  1399. .1704
                    /rpt_family="MaLR"
     repeat_region  1932. .2226
                    /rpt_family="Alu"
     repeat_region  2485. .2635
                    /rpt_family="L1"
     repeat_region  4039. .4165
                    /rpt_family="ERVL"
     repeat_region  4324. .4583
                    /rpt_family="ERVL"
     repeat_region  4599. .4703
                    /rpt_family="ERV1"
     repeat_region  8042. .8478
                    /rpt_family="L1"
     repeat_region  8542. .8809
                    /rpt_family="Alu"
     repeat_region  8879. .8941
                    /rpt_family="MIR"
     repeat_region  9101. .9145
                    /rpt_family="(TAA)n"
     repeat_region  9221. .9262
                    /rpt_family="AT_rich"
     repeat_region  9268. .9476
                    /rpt_family="MIR"
     repeat_region  9759. .11034
                    /rpt_family="ERVL"
```

```
repeat_region   11142. .11311
                /rpt_family="ERVL"
repeat_region   11722. .12119
                /rpt_family="ERVL"
repeat_region   12345. .12454
                /rpt_family="MER2_type"
repeat_region   12455. .12838
                /rpt_family="MaLR"
repeat_region   12839. .13227
                /rpt_family="MER2_type"
repeat_region   13618. .13846
                /rpt_family="MIR"
repeat_region   14052. .14340
                /rpt_family="Alu"
repeat_region   14449. .14502
                /rpt_family="MIR"
repeat_region   15587. .15641
                /rpt_family="(TG)n"
repeat_region   15657. .15738
                /rpt_family="MIR"
repeat_region   15899. .16020
                /rpt_family="L2"
repeat_region   22241. .22595
                /rpt_family="MaLR"
repeat_region   22597. .22653
                /rpt_family="(CA)n"
repeat_region   23206. .23366
                /rpt_family="L2"
repeat_region   23401. .23773
                /rpt_family="MaLR"
repeat_region   27341. .27627
                /rpt_family="L1"
repeat_region   27628. .27927
                /rpt_family="Alu"
repeat_region   27928. .29009
                /rpt_family="L1"
repeat_region   29010. .29363
                /rpt_family="MaLR"
repeat_region   30946. .31295
                /rpt_family="MaLR"
repeat_region   31296. .32186
                /rpt_family="L1"
repeat_region   32187. .32524
                /rpt_family="L1"
repeat_region   32741. .33049
                /rpt_family="L1"
repeat_region   33050. .33152
                /rpt_family="L1"
repeat_region   33136. .33443
                /rpt_family="L1"
repeat_region   33510. .33738
                /rpt_family="MIR"
repeat_region   35210. .35327
                /rpt_family="MIR"
repeat_region   35787. .35814
                /rpt_family="AT_rich"
repeat_region   37663. .37694
                /rpt_family="AT_rich"
repeat_region   38917. .39027
                /rpt_family="L2"
repeat_region   41453. .41689
                /rpt_family="L2"
repeat_region   41776. .41820
                /rpt_family="MIR"

  Query Match             89.5%;  Score 17;  DB 9;  Length 92611;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ATTATTGATTCTGTAGG 18
              |||||||||||||||||
Db      82101 ATTATTGATTCTGTAGG 82085
```

```
RESULT 21
AC091621/c
LOCUS       AC091621               101443 bp    DNA     linear   PRI 04-JAN-2003
DEFINITION  Papio anubis clone RP41-57J19, complete sequence.
ACCESSION   AC091621
VERSION     AC091621.4  GI:27497219
KEYWORDS    HTG.
SOURCE      Papio anubis (olive baboon)
  ORGANISM  Papio anubis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 101443)
  AUTHORS   Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
            Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
            Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
            Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
            Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
            Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
            Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 101443)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2001) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 101443)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JAN-2003) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On Jan 4, 2003 this sequence version replaced gi:18640641.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: cci
                 Center clone name: 057J19

            This sequence was finished as follows unless otherwise noted:
            all regions were double-stranded, sequenced with an
            alternate chemistry, or covered by high quality data
            (i.e., phred quality >= 30); an attempt was made to resolve
            all sequencing problems, such as compressions and repeats;
            all regions were covered by at least one plasmid subclone
            or more than one M13 subclone; and the assembly was confirmed
            by restriction digest.

            CLONE LENGTH: This sequence represents the entire insert of
            this clone unless otherwise noted. If there are overlapping
            clones, the overlaps are noted in the beginning and end of
            the Features section.
FEATURES             Location/Qualifiers
     source          1. .101443
                     /organism="Papio anubis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9555"
                     /clone="RP41-57J19"
                     /clone_lib="RP41"
     misc_feature    <1. .31435
                     /note="this sequence is not the entire insert of clone
                     RP41-57J19; clone overlaps with GenBank Accession Number
                     AC092758 (nucleotides 123789-171690) clone RP41-22J16
                     (center project name ccj); this annotated segment
                     represents overlap with nucleotides 140269-171690 of
                     AC092758"
```

```
     misc_feature    309. .326
                     /note="single clone coverage"
     misc_feature    85858. .>101443
                     /note="this sequence is not the entire insert of clone
                     RP41-57J19; clone overlaps with GenBank Accession Number
                     AC098811 (nucleotides 1-60859) clone RP41-367L14 (center
                     project name csm); this annotated segment represents
                     overlap with nucleotides 1-15586 of AC098811"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 9;  Length 101443;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ATTATTGATTCTGTAGG 18
              |||||||||||||||||
Db     100356 ATTATTGATTCTGTAGG 100340


RESULT 22
LMFLCHR31_02/c
WPCOMMENT
  Sequence split into 22 fragments  LOCUS LMFLCHR31 Accession AL499621
     Fragment Name      Begin     End
     LMFLCHR31_00           1    110000
     LMFLCHR31_01      100001    210000
     LMFLCHR31_02      200001    310000
     LMFLCHR31_03      300001    410000
     LMFLCHR31_04      400001    510000
     LMFLCHR31_05      500001    610000
     LMFLCHR31_06      600001    710000
     LMFLCHR31_07      700001    810000
     LMFLCHR31_08      800001    910000
     LMFLCHR31_09      900001   1010000
     LMFLCHR31_10     1000001   1110000
     LMFLCHR31_11     1100001   1210000
     LMFLCHR31_12     1200001   1310000
     LMFLCHR31_13     1300001   1410000
     LMFLCHR31_14     1400001   1510000
     LMFLCHR31_15     1500001   1610000
     LMFLCHR31_16     1600001   1710000
     LMFLCHR31_17     1700001   1810000
     LMFLCHR31_18     1800001   1910000
     LMFLCHR31_19     1900001   2010000
     LMFLCHR31_20     2000001   2110000
     LMFLCHR31_21     2100001   2117963
Continuation (3 of 22) of LMFLCHR31 from base 200001 (AL499621 Leishmania major chromosc

  Query Match             89.5%;  Score 17;  DB 2;  Length 110000;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      77355 TTATTGATTCTGTAGGT 77339


RESULT 23
AC092456
LOCUS       AC092456               140805 bp    DNA     linear   HTG 06-JUL-2001
DEFINITION  Homo sapiens chromosome CTD clone CTD-2155H16, WORKING DRAFT
            SEQUENCE, 23 unordered pieces.
ACCESSION   AC092456
VERSION     AC092456.1  GI:14625102
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 140805)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 140805)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT
            ------------------ Genome Center -------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information ----------------
            Center project name: H_MS2155H16
            -----------------------------------------------------
            ------------- Summary Statistics  -------------
            Sequencing vector: M13; 74%
            Sequencing vector: plasmid; 21%
            Chemistry: Dye-primer ET; 74% of reads
            Chemistry: Dye-terminator Big Dye; 21% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 125167 bases at least Q40
            Consensus quality: 130374 bases at least Q30
            Consensus quality: 134105 bases at least Q20
            Insert size: 106000; agarose-fp
            Insert size: 139359; sum-of-contigs
            Quality coverage: 4.05 in Q20 bases; agarose-fp
            Quality coverage: 3.14 in Q20 bases; sum-of-contigs
            -----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 23 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1975: contig of 1975 bp in length
            *     1976     2075: gap of unknown length
            *     2076     3541: contig of 1466 bp in length
            *     3542     3641: gap of unknown length
            *     3642     4805: contig of 1164 bp in length
            *     4806     4905: gap of unknown length
            *     4906     6516: contig of 1611 bp in length
            *     6517     6616: gap of unknown length
            *     6617     8152: contig of 1536 bp in length
            *     8153     8252: gap of unknown length
            *     8253    11125: contig of 2873 bp in length
            *    11126    11225: gap of unknown length
            *    11226    13188: contig of 1963 bp in length
            *    13189    13288: gap of unknown length
            *    13289    17124: contig of 3836 bp in length
            *    17125    17224: gap of unknown length
            *    17225    19911: contig of 2687 bp in length
            *    19912    20011: gap of unknown length
            *    20012    23136: contig of 3125 bp in length
            *    23137    23236: gap of unknown length
            *    23237    28202: contig of 4966 bp in length
            *    28203    28302: gap of unknown length
            *    28303    33141: contig of 4839 bp in length
            *    33142    33241: gap of unknown length
            *    33242    37167: contig of 3926 bp in length
            *    37168    37267: gap of unknown length
            *    37268    41815: contig of 4548 bp in length
            *    41816    41915: gap of unknown length
            *    41916    46824: contig of 4909 bp in length
            *    46825    46924: gap of unknown length
            *    46925    52358: contig of 5434 bp in length
            *    52359    52458: gap of unknown length
            *    52459    58556: contig of 6098 bp in length
            *    58557    58656: gap of unknown length
            *    58657    65371: contig of 6715 bp in length
            *    65372    65471: gap of unknown length
            *    65472    72114: contig of 6643 bp in length
```

```
          *    72115    72214: gap of unknown length
          *    72215    81478: contig of 9264 bp in length
          *    81479    81578: gap of unknown length
          *    81579    90420: contig of 8842 bp in length
          *    90421    90520: gap of unknown length
          *    90521   113878: contig of 23358 bp in length
          *   113879   113978: gap of unknown length
          *   113979   140805: contig of 26827 bp in length.
FEATURES             Location/Qualifiers
     source          1. .140805
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="CTD"
                     /clone="CTD-2155H16"
     misc_feature    1. .1975
                     /note="assembly_name:Contig27"
     misc_feature    2076. .3541
                     /note="assembly_name:Contig28"
     misc_feature    3642. .4805
                     /note="assembly_name:Contig29"
     misc_feature    4906. .6516
                     /note="assembly_name:Contig32"
     misc_feature    6617. .8152
                     /note="assembly_name:Contig33"
     misc_feature    8253. .11125
                     /note="assembly_name:Contig34"
     misc_feature    11226. .13188
                     /note="assembly_name:Contig35"
     misc_feature    13289. .17124
                     /note="assembly_name:Contig37"
     misc_feature    17225. .19911
                     /note="assembly_name:Contig38"
     misc_feature    20012. .23136
                     /note="assembly_name:Contig39"
     misc_feature    23237. .28202
                     /note="assembly_name:Contig40"
     misc_feature    28303. .33141
                     /note="assembly_name:Contig41"
     misc_feature    33242. .37167
                     /note="assembly_name:Contig42"
     misc_feature    37268. .41815
                     /note="assembly_name:Contig43"
     misc_feature    41916. .46824
                     /note="assembly_name:Contig44"
     misc_feature    46925. .52358
                     /note="assembly_name:Contig45"
     misc_feature    52459. .58556
                     /note="assembly_name:Contig46"
     misc_feature    58657. .65371
                     /note="assembly_name:Contig47"
     misc_feature    65472. .72114
                     /note="assembly_name:Contig48"
     misc_feature    72215. .81478
                     /note="assembly_name:Contig49"
     misc_feature    81579. .90420
                     /note="assembly_name:Contig50"
     misc_feature    90521. .113878
                     /note="assembly_name:Contig51"
     misc_feature    113979. .140805
                     /note="assembly_name:Contig52"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 140805;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db      24869 AATTATTGATTCTGTAG 24885


RESULT 24
```

```
AC007670
LOCUS       AC007670               150070 bp    DNA     linear   HTG 24-MAR-2000
DEFINITION  Homo sapiens clone RP11-14A7, WORKING DRAFT SEQUENCE, 11 unordered
            pieces.
ACCESSION   AC007670
VERSION     AC007670.2  GI:7321636
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 150070)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-14A7
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150070)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
            Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
            Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
            Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 24, 2000 this sequence version replaced gi:4914394.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L783
                Center clone name: 14_A_7
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-primer-amersham; 85% of reads
                Chemistry: Dye-terminator Big Dye; 15% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 141263 bases at least Q40
                Consensus quality: 145915 bases at least Q30
                Consensus quality: 147932 bases at least Q20
                Insert size: 168000; agarose-fp
                Insert size: 149070; sum-of-contigs
                Quality coverage: 4.4 in Q20 bases; agarose-fp
                Quality coverage: 4.9 in Q20 bases.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1089: contig of 1089 bp in length
            *     1090     1189: gap of 100 bp
            *     1190     2281: contig of 1092 bp in length
            *     2282     2381: gap of 100 bp
            *     2382     6681: contig of 4300 bp in length
            *     6682     6781: gap of 100 bp
            *     6782    11902: contig of 5121 bp in length
```

```
            *    11903    12002: gap of 100 bp
            *    12003    19315: contig of 7313 bp in length
            *    19316    19415: gap of 100 bp
            *    19416    27440: contig of 8025 bp in length
            *    27441    27540: gap of 100 bp
            *    27541    36885: contig of 9345 bp in length
            *    36886    36985: gap of 100 bp
            *    36986    51343: contig of 14358 bp in length
            *    51344    51443: gap of 100 bp
            *    51444    65766: contig of 14323 bp in length
            *    65767    65866: gap of 100 bp
            *    65867    84896: contig of 19030 bp in length
            *    84897    84996: gap of 100 bp
            *    84997   150070: contig of 65074 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150070
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-14A7"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1089
                     /note="assembly_fragment"
     misc_feature    1190. .2281
                     /note="assembly_fragment"
     misc_feature    2382. .6681
                     /note="assembly_fragment"
     misc_feature    6782. .11902
                     /note="assembly_fragment"
     misc_feature    12003. .19315
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     misc_feature    19416. .27440
                     /note="assembly_fragment"
     misc_feature    27541. .36885
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    36986. .51343
                     /note="assembly_fragment"
     misc_feature    51444. .65766
                     /note="assembly_fragment"
     misc_feature    65867. .84896
                     /note="assembly_fragment"
     misc_feature    84997. .150070
                     /note="assembly_fragment"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 150070;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db      48335 AATTATTGATTCTGTAG 48351


RESULT 25
AC091661/c
LOCUS       AC091661              153270 bp    DNA     linear   HTG 03-MAY-2002
DEFINITION  Pan troglodytes clone RP43-168A12, WORKING DRAFT SEQUENCE, 2
            ordered pieces.
ACCESSION   AC091661
VERSION     AC091661.2  GI:20429317
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 153270)
  AUTHORS   Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
            Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
            Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
            Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
            Margulies,E.H., Masiello,C., Maskeri,B., Mastrian,S.D.,
            McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
            Portnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,
            Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 153270)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2001) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 153270)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2002) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On May 3, 2002 this sequence version replaced gi:14042998.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: ceh
                 Center clone name: 168A12

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 152627 bases at least Q40
                 Consensus quality: 152987 bases at least Q30
                 Consensus quality: 153127 bases at least Q20
                 Insert size: 123000; agarose-fp
                 Insert size: 153170; sum-of-contigs
                 Quality coverage: 13.06x in Q20 bases; agarose-fp
                 Quality coverage: 10.49x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   103005: contig of 103005 bp in length
            *  103006   103105: gap of unknown length
            *  103106   153270: contig of 50165 bp in length.
FEATURES             Location/Qualifiers
     source          1. .153270
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-168A12"
                     /clone_lib="RP43"
```

```
     misc_feature    1. .103005
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     misc_feature    1. .4283
                     /note="clone overlaps with GenBank Accession Number
                     AC091718 clone RP43-184N18 (center project name ceg)"
     misc_feature    78757. .153270
                     /note="clone overlaps with GenBank Accession Number
                     AC096879 clone RP43-179H6 (center project name clm)"
     misc_feature    103106. .153270
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 153270;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ATTATTGATTCTGTAGG 18
              |||||||||||||||||
Db      50279 ATTATTGATTCTGTAGG 50263


RESULT 26
AC098817/c
LOCUS       AC098817               154436 bp    DNA     linear   PRI 21-FEB-2002
DEFINITION  Homo sapiens BAC clone RP11-92A1 from 2, complete sequence.
ACCESSION   AC098817 AC019324
VERSION     AC098817.3  GI:18464265
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 154436)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 154436)
  AUTHORS   Isak,A., Cotton,M. and Boyer,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-92A1
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 154436)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 154436)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 154436)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Feb 1, 2002 this sequence version replaced gi:17647046.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0092A01
                Drafting Center: WIBR
            --------------.

NOTICE:  This sequence may not represent the entire insert of this
clone.  It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR:  pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-345F13; the clone sequenced
to the right is RP11-1290B4.  Actual start of this clone is at base
position 1 of RP11-92A1; actual end is at base position 154436 of
RP11-92A1.

Data from AC092456 and AC007670 was used to finish this clone,
AC098817.

The sequence of AC019324 has been incorporated into AC098817.
FEATURES             Location/Qualifiers
     source          1. .154436
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-92A1"
                     /clone_lib="RPCI-11"
     repeat_region   630. .1089
                     /rpt_family="ERV1"
     repeat_region   1295. .1842
                     /rpt_family="ERV1"
     repeat_region   1864. .1919
                     /rpt_family="MaLR"
     repeat_region   1920. .1950
                     /rpt_family="(TTTG)n"
     repeat_region   2224. .2606
                     /rpt_family="MaLR"
     repeat_region   2747. .3376
                     /rpt_family="ERV1"
     repeat_region   3583. .3639
                     /rpt_family="AT_rich"
     repeat_region   4256. .4319
                     /rpt_family="(TAGA)n"
     repeat_region   5742. .6544
                     /rpt_family="L1"
     repeat_region   6552. .6915
                     /rpt_family="L1"
     repeat_region   6921. .8756
                     /rpt_family="L1"
     repeat_region   8756. .9240
                     /rpt_family="L1"
```

```
repeat_region   9241. .9310
                /rpt_family="L1"
repeat_region   9311. .10462
                /rpt_family="L1"
repeat_region   9542. .9563
                /rpt_family="AT_rich"
repeat_region   10466. .11792
                /rpt_family="L1"
repeat_region   11770. .11811
                /rpt_family="AT_rich"
repeat_region   11925. .12123
                /rpt_family="MIR"
repeat_region   12234. .12266
                /rpt_family="(TG)n"
repeat_region   12295. .12871
                /rpt_family="L1"
repeat_region   13506. .13545
                /rpt_family="AT_rich"
repeat_region   13683. .13988
                /rpt_family="Alu"
repeat_region   14427. .14533
                /rpt_family="MaLR"
repeat_region   15067. .15094
                /rpt_family="AT_rich"
repeat_region   15544. .15692
                /rpt_family="L1"
misc_feature    18589. .18618
                /note="similar to Homo sapiens EST BG877852
                (NID:g14254942)"
misc_feature    19265. .19715
                /note="similar to Homo sapiens EST BF768765
                (NID:g12116665)"
misc_feature    19308. .19356
                /note="similar to Homo sapiens EST BE879309
                (NID:g10328085)"
misc_feature    19308. .19356
                /note="similar to Homo sapiens EST BG877217
                (NID:g14254307)"
misc_feature    19423. .19778
                /note="similar to Homo sapiens EST BF769298
                (NID:g12117198)"
misc_feature    19442. .19775
                /note="similar to Homo sapiens EST BF769297
                (NID:g12117197)"
misc_feature    19449. .19514
                /note="similar to Homo sapiens EST BE879309
                (NID:g10328085)"
misc_feature    19492. .19514
                /note="similar to Homo sapiens EST BG877217
                (NID:g14254307)"
misc_feature    19702. .19908
                /note="similar to Homo sapiens EST AA577014 (NID:g2354488)
                nm84f02.s1"
misc_feature    19702. .19793
                /note="similar to Homo sapiens EST AL044154
                (NID:g5432379)"
misc_feature    19715. .19793
                /note="similar to Homo sapiens EST BG877519
                (NID:g14254609)"
misc_feature    19717. .19784
                /note="similar to Homo sapiens EST BE879309
                (NID:g10328085)"
misc_feature    19717. .19778
                /note="similar to Homo sapiens EST BG877217
                (NID:g14254307)"
misc_feature    19894. .20466
                /note="similar to Bos taurus EST BI898628 (NID:g16204051)"
misc_feature    20624. .20766
                /note="similar to Homo sapiens EST AW867815
                (NID:g8001776)"
misc_feature    20807. .20867
                /note="similar to Homo sapiens EST AW867815
                (NID:g8001776)"
misc_feature    21020. .21278
                /note="similar to Homo sapiens EST AI377331 (NID:g4187184)
                te60c10.x1"
misc_feature    21376. .21952
                /note="similar to EST BM241349 (NID:g17876619)"
misc_feature    21435. .22319
                /note="similar to Mus musculus EST BI692638
                (NID:g15655267)"

  Query Match             89.5%;  Score 17;  DB 9;  Length 154436;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db      95117 AATTATTGATTCTGTAG 95101


RESULT 27
AL357074
LOCUS       AL357074               157800 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 6 clone RP11-600C19, 6 unordered pieces.
ACCESSION   AL357074
VERSION     AL357074.3  GI:9863808
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burton,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 21, 2000 this sequence version replaced gi:9214057.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA600C19
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 154383 bases at least Q40
            Consensus quality: 155475 bases at least Q30
            Consensus quality: 156261 bases at least Q20
            Insert size: 157300; sum-of-contigs
            Insert size: 150139; 8.4% error; agarose-fp
            Quality coverage: 4.87x in Q20 bases; sum-of-contigs Quality
            coverage: 5.20x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    14793: contig of 14793 bp in length
            *    14794    14893: gap of 100 bp
            *    14894    55692: contig of 40799 bp in length
            *    55693    55792: gap of 100 bp
            *    55793    75176: contig of 19384 bp in length
            *    75177    75276: gap of 100 bp
            *    75277   123901: contig of 48625 bp in length
            *   123902   124001: gap of 100 bp
            *   124002   135576: contig of 11575 bp in length
            *   135577   135676: gap of 100 bp
```

```
           *   135677   157800: contig of 22124 bp in length.
FEATURES             Location/Qualifiers
     source          1. .157800
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-600C19"
                     /clone_lib="RPCI-11.3"
     misc_feature    1. .14793
                     /note="assembly_fragment:01602
                     fragment_chain:1
                     clone_end:T7
                     vector_side:left"
     misc_feature    14894. .55692
                     /note="assembly_fragment:00080
                     fragment_chain:1"
     misc_feature    55793. .75176
                     /note="assembly_fragment:00503
                     fragment_chain:2"
     misc_feature    75277. .123901
                     /note="assembly_fragment:01560
                     fragment_chain:2"
     misc_feature    124002. .135576
                     /note="assembly_fragment:00932
                     fragment_chain:3"
     misc_feature    135677. .157800
                     /note="assembly_fragment:01303
                     fragment_chain:3"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 157800;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db     106015 TTATTGATTCTGTAGGT 106031


RESULT 28
AC013613/c
LOCUS       AC013613               159790 bp    DNA     linear   HTG 10-SEP-2000
DEFINITION  Homo sapiens clone RP11-10M11, WORKING DRAFT SEQUENCE, 11 unordered
            pieces.
ACCESSION   AC013613
VERSION     AC013613.4  GI:10047743
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 159790)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-10M11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 159790)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 10, 2000 this sequence version replaced gi:7230059.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3104
                Center clone name: 10_M_11
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 152125 bases at least Q40
                Consensus quality: 156165 bases at least Q30
                Consensus quality: 157620 bases at least Q20
                Insert size: 163000; agarose-fp
                Insert size: 158790; sum-of-contigs
                Quality coverage: 4.6 in Q20 bases; agarose-fp
                Quality coverage: 4.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     6745: contig of 6745 bp in length
            *     6746     6845: gap of 100 bp
            *     6846    10699: contig of 3854 bp in length
            *    10700    10799: gap of 100 bp
            *    10800    34927: contig of 24128 bp in length
            *    34928    35027: gap of 100 bp
            *    35028    44799: contig of 9772 bp in length
            *    44800    44899: gap of 100 bp
            *    44900    54273: contig of 9374 bp in length
            *    54274    54373: gap of 100 bp
            *    54374    64630: contig of 10257 bp in length
            *    64631    64730: gap of 100 bp
            *    64731    78802: contig of 14072 bp in length
            *    78803    78902: gap of 100 bp
            *    78903    99071: contig of 20169 bp in length
            *    99072    99171: gap of 100 bp
            *    99172   127231: contig of 28060 bp in length
            *   127232   127331: gap of 100 bp
            *   127332   152287: contig of 24956 bp in length
            *   152288   152387: gap of 100 bp
            *   152388   159790: contig of 7403 bp in length.
FEATURES             Location/Qualifiers
     source          1. .159790
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-10M11"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .6745
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    6846. .10699
                     /note="assembly_fragment"
     misc_feature    10800. .34927
                     /note="assembly_fragment"
     misc_feature    35028. .44799
                     /note="assembly_fragment"
     misc_feature    44900. .54273
                     /note="assembly_fragment"
```

```
     misc_feature    54374. .64630
                     /note="assembly_fragment"
     misc_feature    64731. .78802
                     /note="assembly_fragment"
     misc_feature    78903. .99071
                     /note="assembly_fragment"
     misc_feature    99172. .127231
                     /note="assembly_fragment"
     misc_feature    127332. .152287
                     /note="assembly_fragment"
     misc_feature    152388. .159790
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 159790;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db     119355 TTATTGATTCTGTAGGT 119339


RESULT 29
AC099340/c
LOCUS       AC099340               164958 bp    DNA     linear   PRI 29-MAY-2002
DEFINITION  Homo sapiens BAC clone RP11-214O14 from 4, complete sequence.
ACCESSION   AC099340 AC024369
VERSION     AC099340.3  GI:20429590
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164958)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 164958)
  AUTHORS   Boatright,E., Kozlowicz,A. and Boyer,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-214O14
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 164958)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 164958)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 164958)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On May 3, 2002 this sequence version replaced gi:17386499.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0214O14
                Drafting Center: WIBR
            --------------.
```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-398H1; the clone sequenced to the right is RP11-307L18. Actual start of this clone is at base position 1 of RP11-214O14; actual end is at base position 164958 of RP11-214O14.

Polymorphisms exist between AC108045 and AC099340. Data from AC107053 and AC108045 was used to finish AC099340.

The sequence of AC024369 has been incorporated into AC099340.

```
FEATURES             Location/Qualifiers
     source          1. .164958
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4"
                     /clone="RP11-214O14"
                     /clone_lib="RPCI-11"
     repeat_region   114. .430
                     /rpt_family="L1"
     repeat_region   441. .529
                     /rpt_family="Alu"
     repeat_region   530. .682
                     /rpt_family="L1"
     repeat_region   683. .998
                     /rpt_family="Alu"
     repeat_region   999. .1069
                     /rpt_family="L1"
     repeat_region   1070. .1106
                     /rpt_family="AT_rich"
     repeat_region   1735. .1787
                     /rpt_family="MIR"
     repeat_region   1791. .2025
                     /rpt_family="L1"
     repeat_region   2034. .2347
                     /rpt_family="Alu"
     repeat_region   2384. .2533
                     /rpt_family="L1"
     repeat_region   2535. .2666
                     /rpt_family="Alu"
     repeat_region   2668. .3178
                     /rpt_family="ERV1"
```

```
repeat_region   3181. .3543
                /rpt_family="L1"
repeat_region   3544. .3774
                /rpt_family="L1"
repeat_region   3752. .3898
                /rpt_family="L1"
repeat_region   3908. .4293
                /rpt_family="MER2_type"
repeat_region   4330. .4636
                /rpt_family="Alu"
repeat_region   4751. .5062
                /rpt_family="Alu"
repeat_region   5268. .5577
                /rpt_family="Alu"
repeat_region   5685. .5915
                /rpt_family="Alu"
repeat_region   5916. .6071
                /rpt_family="L1"
repeat_region   6093. .6183
                /rpt_family="(TCCC)n"
repeat_region   6205. .6369
                /rpt_family="Alu"
repeat_region   6373. .6618
                /rpt_family="L1"
repeat_region   6709. .7002
                /rpt_family="Alu"
repeat_region   7003. .7070
                /rpt_family="(TA)n"
repeat_region   7076. .7310
                /rpt_family="L1"
repeat_region   7328. .7558
                /rpt_family="L1"
repeat_region   7559. .7835
                /rpt_family="Alu"
repeat_region   7836. .8172
                /rpt_family="L1"
repeat_region   8180. .8460
                /rpt_family="Alu"
repeat_region   8836. .9160
                /rpt_family="MaLR"
repeat_region   9161. .9663
                /rpt_family="L2"
repeat_region   10587. .10742
                /rpt_family="MIR"
repeat_region   11088. .11226
                /rpt_family="Alu"
repeat_region   12985. .13139
                /rpt_family="MIR"
repeat_region   13977. .14228
                /rpt_family="L1"
repeat_region   14229. .14525
                /rpt_family="Alu"
repeat_region   14526. .14941
                /rpt_family="L1"
repeat_region   14925. .15420
                /rpt_family="L1"
repeat_region   15421. .15766
                /rpt_family="Alu"
repeat_region   15767. .16239
                /rpt_family="L1"
repeat_region   16240. .18114
                /rpt_family="L1"
repeat_region   18117. .18219
                /rpt_family="Alu"
repeat_region   18230. .18444
                /rpt_family="L2"
repeat_region   18668. .19061
                /rpt_family="MER2_type"
repeat_region   19062. .19362
                /rpt_family="Alu"
repeat_region   19363. .19445
                /rpt_family="MER2_type"
repeat_region   19581. .19912
                /rpt_family="Alu"
repeat_region   19924. .20305
                /rpt_family="Alu"
repeat_region   20329. .20926
                /rpt_family="L1"
repeat_region   20927. .21239
                /rpt_family="Alu"
repeat_region   21240. .21267
```

Query Match 89.5%; Score 17; DB 9; Length 164958;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      70971 TTATTGATTCTGTAGGT 70955
```

```
RESULT 30
AC119280/c
LOCUS       AC119280              172263 bp    DNA     linear   HTG 09-MAR-2003
DEFINITION  Mus musculus clone RP24-462F23, WORKING DRAFT SEQUENCE, 6 unordered
            pieces.
ACCESSION   AC119280
VERSION     AC119280.3  GI:28893724
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 172263)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP24-462F23
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172263)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 172263)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
```

```
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 9, 2003 this sequence version replaced gi:28626749.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L25706
                Center clone name: 462_F_23
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 170828 bases at least Q40
                Consensus quality: 171526 bases at least Q30
                Consensus quality: 171702 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 171763; sum-of-contigs
                Quality coverage: 9.9 in Q20 bases; agarose-fp
                Quality coverage: 9.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1342: contig of 1342 bp in length
            *     1343     1442: gap of 100 bp
            *     1443     2735: contig of 1293 bp in length
            *     2736     2835: gap of 100 bp
            *     2836     7640: contig of 4805 bp in length
            *     7641     7740: gap of 100 bp
            *     7741    18392: contig of 10652 bp in length
            *    18393    18492: gap of 100 bp
            *    18493    65941: contig of 47449 bp in length
            *    65942    66041: gap of 100 bp
            *    66042   172263: contig of 106222 bp in length.
FEATURES             Location/Qualifiers
     source          1. .172263
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP24-462F23"
                     /clone_lib="RPCI-24 Male Mouse BAC"
     misc_feature    1. .1342
                     /note="assembly_fragment"
     misc_feature    1443. .2735
                     /note="assembly_fragment"
     misc_feature    2836. .7640
                     /note="assembly_fragment"
     misc_feature    7741. .18392
                     /note="assembly_fragment"
     misc_feature    18493. .65941
                     /note="assembly_fragment"
     misc_feature    66042. .172263
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 172263;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db      57734 AATTATTGATTCTGTAG 57718


RESULT 31
AC098811/c
LOCUS       AC098811              182644 bp    DNA     linear   PRI 28-JAN-2003
DEFINITION  Papio anubis clone RP41-367L14, complete sequence.
ACCESSION   AC098811
VERSION     AC098811.3  GI:21903573
KEYWORDS    HTG.
SOURCE      Papio anubis (olive baboon)
  ORGANISM  Papio anubis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 182644)
  AUTHORS   Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
            Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
            Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
            Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
            Margulies,E.H., Masiello,C., Maskeri,B., Mastrian,S.D.,
            McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
            Portnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,
            Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182644)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-2001) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 182644)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   4  (bases 1 to 182644)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2003) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On Jul 18, 2002 this sequence version replaced gi:18642737.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: csm
                 Center clone name: 367L14

            This sequence was finished as follows unless otherwise noted:
            all regions were double-stranded, sequenced with an
            alternate chemistry, or covered by high quality data
            (i.e., phred quality >= 30); an attempt was made to resolve
            all sequencing problems, such as compressions and repeats;
            all regions were covered by at least one plasmid subclone
            or more than one M13 subclone; and the assembly was confirmed
            by restriction digest.

            CLONE LENGTH: This sequence represents the entire insert of
```

this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features section.

```
FEATURES             Location/Qualifiers
     source          1. .182644
                     /organism="Papio anubis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9555"
                     /clone="RP41-367L14"
                     /clone_lib="RP41"
     misc_feature    16990. .17127
                     /note="single clone coverage"
     misc_feature    18401. .18470
                     /note="single clone coverage"
     misc_feature    18709. .18756
                     /note="single clone coverage"
     misc_feature    48224. .48240
                     /note="single clone coverage"
     misc_feature    75317. .75463
                     /note="single clone coverage"
     misc_feature    114474. .114497
                     /note="single clone coverage"
     misc_feature    143993. .144076
                     /note="single clone coverage"
ORIGIN

  Query Match          89.5%;  Score 17;  DB 9;  Length 182644;
  Best Local Similarity  100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ATTATTGATTCTGTAGG 18
              |||||||||||||||||
Db      14499 ATTATTGATTCTGTAGG 14483
```

```
RESULT 32
AC102150
LOCUS       AC102150               183774 bp    DNA     linear   HTG 29-SEP-2004
DEFINITION  Mus musculus chromosome 7 clone RP23-310N5 map 7, *** SEQUENCING IN
            PROGRESS ***, 2 ordered pieces.
ACCESSION   AC102150
VERSION     AC102150.9  GI:52839788
KEYWORDS    HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 183774)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 7, clone RP23-310N5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 183774)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 183774)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-SEP-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 29, 2004 this sequence version replaced gi:52782635.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            -------------- Project Information
                Center project name: L18115
                Center clone name: 310_N_5
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    72244: contig of 72244 bp in length
            *    72245    72344: gap of unknown length
            *    72345   183774: contig of 111430 bp in length.
FEATURES             Location/Qualifiers
     source          1. .183774
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="7"
                     /map="7"
                     /clone="RP23-310N5"
                     /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN

  Query Match          89.5%;  Score 17;  DB 2;  Length 183774;
  Best Local Similarity  100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      50297 TTATTGATTCTGTAGGT 50313
```

```
RESULT 33
AC133083
LOCUS       AC133083              194648 bp    DNA     linear   HTG 17-AUG-2004
DEFINITION  Mus musculus chromosome 12 clone RP24-342G5, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC133083
VERSION     AC133083.2  GI:51315649
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 194648)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 194648)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-SEP-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 194648)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Aug 17, 2004 this sequence version replaced gi:22748525.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            ---------- Project Information ----------
            Center project name: M_BB0342G05
            -----------------------------------------------------
            ---------- Summary Statistics  ----------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 192105 bases at least Q40
            Consensus quality: 192564 bases at least Q30
            Consensus quality: 192894 bases at least Q20
            -----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1097: contig of 1097 bp in length
            *     1098     1197: gap of unknown length
            *     1198     2722: contig of 1525 bp in length
            *     2723     2822: gap of unknown length
            *     2823     4221: contig of 1399 bp in length
            *     4222     4321: gap of unknown length
            *     4322    23429: contig of 19108 bp in length
            *    23430    23529: gap of unknown length
            *    23530   109776: contig of 86247 bp in length
            *   109777   109876: gap of unknown length
            *   109877   194648: contig of 84772 bp in length.
FEATURES             Location/Qualifiers
     source          1. .194648
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="12"
                     /clone="RP24-342G5"
     misc_feature    1. .1097
                     /note="assembly_name:Contig4"
     misc_feature    1198. .2722
                     /note="assembly_name:Contig5"
     misc_feature    2823. .4221
                     /note="assembly_name:Contig6"
     misc_feature    4322. .23429
                     /note="assembly_name:Contig7"
     misc_feature    23530. .109776
                     /note="assembly_name:Contig8
                     clone_end:T7
                     vector_side:right"
     misc_feature    109877. .194648
                     /note="assembly_name:Contig9"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 194648;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      95487 TTATTGATTCTGTAGGT 95503


RESULT 34
AL929408/c
LOCUS       AL929408              196915 bp    DNA     linear   ROD 21-NOV-2002
DEFINITION  Mouse DNA sequence from clone RP23-456B2 on chromosome 4, complete
            sequence.
ACCESSION   AL929408
VERSION     AL929408.3  GI:25136793
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 196915)
  AUTHORS   Howden,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 19, 2002 this sequence version replaced gi:24940059.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-456B2 is
```

```
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .196915
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="4"
                     /clone="RP23-456B2"
                     /clone_lib="RPCI-23"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 10;  Length 196915;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db     138493 AATTATTGATTCTGTAG 138477


RESULT 35
AC119027
LOCUS       AC119027               199373 bp    DNA     linear   HTG 12-OCT-2002
DEFINITION  Rattus norvegicus clone CH230-457H17, *** SEQUENCING IN PROGRESS
            ***, 3 unordered pieces.
ACCESSION   AC119027
VERSION     AC119027.5  GI:23908210
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 199373)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199373)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 199373)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Oct 12, 2002 this sequence version replaced gi:21747329.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: GVXM
                Center clone name: CH230-457H17
            ---------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 167882 bases at least Q40
                Consensus quality: 170782 bases at least Q30
                Consensus quality: 172278 bases at least Q20
                Estimated insert size: 178998; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   195718: contig of 195718 bp in length
            *   195719   195818: gap of unknown length
            *   195819   197296: contig of 1478 bp in length
            *   197297   197396: gap of unknown length
            *   197397   199373: contig of 1977 bp in length.
FEATURES             Location/Qualifiers
     source          1. .199373
```

```
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-457H17"
     misc_feature    1. .1418
                     /note="wgs_contig"
     misc_feature    14276. .15866
                     /note="wgs_contig"
     misc_feature    complement(187609. .188528)
                     /note="clone_boundary
                     clone_end:T7
                     site:MboI
                     end_sequence:RXBMC45TJ"
     misc_feature    189753. .191324
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    193193. .195718
                     /note="wgs_end_extension
                     clone_end:T7"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 199373;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      75541 TTATTGATTCTGTAGGT 75557


RESULT 36
AC138676
LOCUS       AC138676               204308 bp    DNA     linear   HTG 08-OCT-2004
DEFINITION  Mus musculus chromosome 3 clone RP23-381I16 map 3, *** SEQUENCING
            IN PROGRESS ***, 3 unordered pieces.
ACCESSION   AC138676
VERSION     AC138676.5  GI:53984567
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 204308)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 3, clone RP23-381I16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204308)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
            Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
            Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 204308)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Oct 8, 2004 this sequence version replaced gi:51699720.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            -------------- Project Information
                Center project name: L28926
                Center clone name: 381_I_16
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      945: contig of 945 bp in length
            *      946     1045: gap of unknown length
            *     1046    62113: contig of 61068 bp in length
            *    62114    62213: gap of unknown length
            *    62214   204308: contig of 142095 bp in length.
FEATURES            Location/Qualifiers
     source         1. .204308
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10090"
                    /chromosome="3"
                    /map="3"
                    /clone="RP23-381I16"
                    /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 204308;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      56464 TTATTGATTCTGTAGGT 56480


RESULT 37
AC148095/c
LOCUS       AC148095               210358 bp    DNA     linear   HTG 02-FEB-2004
DEFINITION  Mus musculus chromosome 15 clone RP23-36C9 map 15, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC148095
VERSION     AC148095.1  GI:41581467
```

```
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 210358)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 15, clone RP23-36C9
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 210358)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L30113
                Center clone name: 36_C_9
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 208049 bases at least Q40
                Consensus quality: 209013 bases at least Q30
                Consensus quality: 209426 bases at least Q20
                Insert size: 222000; agarose-fp
                Insert size: 209858; sum-of-contigs
                Quality coverage: 9.1 in Q20 bases; agarose-fp
                Quality coverage: 9.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      684: contig of 684 bp in length
            *      685      784: gap of 100 bp
            *      785     1598: contig of 814 bp in length
            *     1599     1698: gap of 100 bp
            *     1699    18422: contig of 16724 bp in length
            *    18423    18522: gap of 100 bp
            *    18523    50410: contig of 31888 bp in length
            *    50411    50510: gap of 100 bp
            *    50511   100097: contig of 49587 bp in length
            *   100098   100197: gap of 100 bp
            *   100198   210358: contig of 110161 bp in length.
FEATURES             Location/Qualifiers
     source          1. .210358
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="15"
                     /map="15"
                     /clone="RP23-36C9"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1. .684
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    785. .1598
                     /note="assembly_fragment"
     misc_feature    1699. .18422
                     /note="assembly_fragment"
     misc_feature    18523. .50410
                     /note="assembly_fragment"
     misc_feature    50511. .100097
                     /note="assembly_fragment"
     misc_feature    100198. .210358
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 210358;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db       7219 TTATTGATTCTGTAGGT 7203


RESULT 38
AC134898/c
LOCUS       AC134898               213348 bp    DNA     linear   ROD 10-JUL-2004
DEFINITION  Mus musculus BAC clone RP23-478H2 from chromosome 14, complete
            sequence.
ACCESSION   AC134898
VERSION     AC134898.4  GI:50200145
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 213348)
  AUTHORS   Hodges,J., Cotton,M., Bielicki,L., Haglund,K., Meyer,R. and
            Haakenson,W.
  TITLE     The sequence of Mus musculus BAC clone RP23-478H2
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 213348)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 213348)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 213348)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2004) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jul 10, 2004 this sequence version replaced gi:47777599.
            -------------- Genome Center
```

```
    Center: Washington University Genome Sequencing Center
    Center code: WUGSC
    Web site: http://genome.wustl.edu
    Contact: submissions@watson.wustl.edu
-------------- Summary Statistics
    Center project name: M_BA0478H02
--------------.

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
and Minako Tateno in the laboratory of Pieter de Jong
(http://www.chori.org) from female C57BL/6J mouse kidney and/or
brain genomic DNA.  The clone and detailed information can be
obtained from Research Genetics, Inc. (http://www.resgen.com) or
Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
FEATURES             Location/Qualifiers
     source          1. .213348
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="14"
                     /map="14"
                     /clone="RP23-478H2"
                     /clone_lib="RPCI-23"
     repeat_region   825. .937
                     /rpt_family="Alu"
     repeat_region   961. .1133
                     /rpt_family="B4"
     repeat_region   1002. .1139
                     /rpt_family="Alu"
     repeat_region   1362. .1499
                     /rpt_family="MaLR"
     repeat_region   1568. .1669
                     /rpt_family="B2"
     repeat_region   1640. .1701
                     /rpt_family="B2"
     repeat_region   1744. .1882
                     /rpt_family="Alu"
     repeat_region   2738. .2868
                     /rpt_family="Alu"
     repeat_region   2875. .3015
                     /rpt_family="Alu"
     repeat_region   3105. .3284
                     /rpt_family="B2"
     repeat_region   3287. .3469
                     /rpt_family="B2"
     repeat_region   4210. .4341
                     /rpt_family="MaLR"
     repeat_region   4410. .4557
                     /rpt_family="Alu"
     misc_feature    4622. .4626
                     /note="Sequence derived from PCR product of project DNA."
     repeat_region   5208. .5479
                     /rpt_family="B4"
     repeat_region   5836. .6047
                     /rpt_family="B2"
     repeat_region   6609. .6762
                     /rpt_family="Alu"
     repeat_region   7013. .7159
                     /rpt_family="B2"
     repeat_region   7183. .7253
                     /rpt_family="Alu"
     repeat_region   7295. .7464
                     /rpt_family="B4"
     repeat_region   7473. .7661
                     /rpt_family="B2"
     repeat_region   8424. .8500
                     /rpt_family="B4"
     repeat_region   8547. .8673
                     /rpt_family="Alu"
     repeat_region   8674. .8708
                     /rpt_family="B4"
     repeat_region   8709. .8849
                     /rpt_family="B4"
     repeat_region   8893. .9008
                     /rpt_family="MaLR"
     repeat_region   9041. .9088
                     /rpt_family="MaLR"
     repeat_region   9089. .9146
                     /rpt_family="tRNA-Ala-GCY_"
     repeat_region   9147. .9206
                     /rpt_family="MaLR"
     repeat_region   9351. .9525
                     /rpt_family="B4"
     repeat_region   9790. .9994
                     /rpt_family="B2"
     repeat_region   9980. .10110
                     /rpt_family="ERVK"
     repeat_region   10111. .10253
                     /rpt_family="Alu"
     repeat_region   10254. .10265
                     /rpt_family="ERVK"
     repeat_region   10290. .10388
                     /rpt_family="B2"
     repeat_region   10609. .10705
                     /rpt_family="MaLR"
     repeat_region   10703. .10762
                     /rpt_family="MaLR"
     repeat_region   10776. .10925
                     /rpt_family="Alu"
     repeat_region   10942. .11012
                     /rpt_family="ID"
     repeat_region   11370. .11492
                     /rpt_family="ERVK"
     unsure          11387. .11390
                     /note="Sequence derived from one plasmid subclone."
     repeat_region   11624. .11732
                     /rpt_family="Alu"
     repeat_region   11771. .12008
                     /rpt_family="MaLR"
     unsure          11978. .12079
                     /note="Sequence derived from one plasmid subclone."
     repeat_region   12233. .12320
                     /rpt_family="Alu"
     repeat_region   12349. .12525
                     /rpt_family="B2"
     tRNA            12355. .12427
                     /product="tRNA-Ser"
                     /note="Likely pseudogene (HMM Sc=38.74 / Sec struct
                     Sc=-16.29)"
     repeat_region   12554. .12657
                     /rpt_family="B2"
     repeat_region   12667. .12861
                     /rpt_family="B2"
     repeat_region   13239. .13378
                     /rpt_family="Alu"
     repeat_region   13598. .13658
```

```
                     /rpt_family="L1"
     repeat_region   13753. .13802
                     /rpt_family="MaLR"
     repeat_region   13784. .13953
                     /rpt_family="B2"
     repeat_region   13953. .14041
                     /rpt_family="B2"
     repeat_region   14034. .14376
                     /rpt_family="MaLR"
     repeat_region   14743. .14855
                     /rpt_family="Alu"
     repeat_region   14771. .14871
                     /rpt_family="MaLR"
     repeat_region   14856. .14883
                     /rpt_family="B4"
     repeat_region   14884. .14940
                     /rpt_family="L1"
     repeat_region   14956. .15019
                     /rpt_family="B2"
     repeat_region   14998. .15927
                     /rpt_family="RMER1B"

  Query Match             89.5%;  Score 17;  DB 10;  Length 213348;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db     130432 AATTATTGATTCTGTAG 130416


RESULT 39
AC098411/c
LOCUS       AC098411               221618 bp    DNA     linear   HTG 13-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-2G8, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC098411
VERSION     AC098411.5  GI:24941356
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 221618)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
    TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 221618)
  AUTHORS   Worley,K.C.
    TITLE   Direct Submission
  JOURNAL   Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 221618)
  AUTHORS   Worley,K.C.
    TITLE   Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 13, 2002 this sequence version replaced gi:23195684.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: TURA
                Center clone name: CH230-2G8
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 209367 bases at least Q40
                Consensus quality: 210730 bases at least Q30
                Consensus quality: 211730 bases at least Q20
                Estimated insert size: 215293; sum-of-contigs estimation
                Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   172380: contig of 172380 bp in length
            *   172381   172480: gap of unknown length
            *   172481   219193: contig of 46713 bp in length
            *   219194   219293: gap of unknown length
            *   219294   220403: contig of 1110 bp in length
            *   220404   220503: gap of unknown length
            *   220504   221618: contig of 1115 bp in length.
FEATURES            Location/Qualifiers
```

```
     source          1. .221618
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-2G8"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 221618;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAG 17
              |||||||||||||||||
Db     170439 AATTATTGATTCTGTAG 170423


RESULT 40
AC103654
LOCUS       AC103654               235315 bp    DNA     linear   HTG 23-MAR-2004
DEFINITION  Mus musculus chromosome 12 clone RP23-259L1 map 12, *** SEQUENCING
            IN PROGRESS ***, 6 unordered pieces.
ACCESSION   AC103654
VERSION     AC103654.4  GI:45642876
KEYWORDS    HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 235315)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 12, clone RP23-259L1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 235315)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 235315)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 23, 2004 this sequence version replaced gi:24182063.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L19529
                Center clone name: 259_L_1
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    50372: contig of 50372 bp in length
            *    50373    50472: gap of 100 bp
            *    50473    55828: contig of 5356 bp in length
            *    55829    55928: gap of 100 bp
            *    55929   114632: contig of 58704 bp in length
            *   114633   114732: gap of 100 bp
            *   114733   116754: contig of 2022 bp in length
            *   116755   116854: gap of 100 bp
            *   116855   163756: contig of 46902 bp in length
            *   163757   163856: gap of 100 bp
            *   163857   235315: contig of 71459 bp in length.
FEATURES             Location/Qualifiers
     source          1. .235315
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="12"
                     /map="12"
                     /clone="RP23-259L1"
                     /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 235315;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db      65661 TTATTGATTCTGTAGGT 65677


RESULT 41
AC125675/c
LOCUS       AC125675               241151 bp    DNA     linear   HTG 19-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-237O17, WORKING DRAFT SEQUENCE, 7
            unordered pieces.
ACCESSION   AC125675
VERSION     AC125675.4  GI:25072983
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
```

Rattus.
REFERENCE 1 (bases 1 to 241151)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 241151)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 241151)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 19, 2002 this sequence version replaced gi:23676845.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GTPS
Center clone name: CH230-237O17
---------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 229311 bases at least Q40
Consensus quality: 231059 bases at least Q30
Consensus quality: 232099 bases at least Q20
Estimated insert size: 228904; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
---
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3625: contig of 3625 bp in length
* 3626 3725: gap of unknown length
* 3726 234785: contig of 231060 bp in length
* 234786 234885: gap of unknown length
* 234886 236114: contig of 1229 bp in length
* 236115 236214: gap of unknown length
* 236215 237456: contig of 1242 bp in length
* 237457 237556: gap of unknown length
* 237557 238736: contig of 1180 bp in length
* 238737 238836: gap of unknown length
* 238837 239890: contig of 1054 bp in length
* 239891 239990: gap of unknown length
* 239991 241151: contig of 1161 bp in length.

FEATURES Location/Qualifiers
source 1. .241151
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-237O17"
misc_feature 1. .1578
/note="wgs_end_extension
clone_end:T7"
misc_feature 3726. .4740
/note="wgs_end_extension
clone_end:T7"
misc_feature complement(5255. .6122)
/note="clone_boundary
clone_end:T7
site:
end_sequence:BZ103403"
misc_feature 227618. .228487
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:BZ103407"

ORIGIN

Query Match 89.5%; Score 17; DB 2; Length 241151;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTAGGT 19

```
                    ||||||||||||||||||
Db      42391 TTATTGATTCTGTAGGT 42375


RESULT 42
AC122943
LOCUS       AC122943               256265 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-3N3, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC122943
VERSION     AC122943.5  GI:30467080
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 256265)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 256265)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 256265)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 9, 2003 this sequence version replaced gi:22772509.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GAOA
                Center clone name: CH230-3N3
            -------------- Summary Statistics
                Assembly program: Atlas;
                Consensus quality: 230047 bases at least Q40
                Consensus quality: 232686 bases at least Q30
                Consensus quality: 234367 bases at least Q20
                Estimated insert size: 240988; sum-of-contigs estimation
                Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   196493: contig of 196493 bp in length
            *   196494   196593: gap of unknown length
            *   196594   253760: contig of 57167 bp in length
            *   253761   253860: gap of unknown length
            *   253861   255090: contig of 1230 bp in length
            *   255091   255190: gap of unknown length
            *   255191   256265: contig of 1075 bp in length.
FEATURES             Location/Qualifiers
     source          1. .256265
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-3N3"
     misc_feature    1. .1416
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    complement(2614. .3386)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH313649"
     misc_feature    9305. .9839
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH313651"
```

```
     misc_feature    196594. .198258
                     /note="wgs_end_extension
                     clone_end:Sp6"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 256265;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db     253253 TTATTGATTCTGTAGGT 253269


RESULT 43
AC110676
LOCUS       AC110676               273331 bp    DNA     linear   HTG 08-OCT-2002
DEFINITION  Rattus norvegicus clone CH230-213N13, *** SEQUENCING IN PROGRESS
            ***, 3 unordered pieces.
ACCESSION   AC110676
VERSION     AC110676.4  GI:23195453
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 273331)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 273331)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 273331)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 19, 2002 this sequence version replaced gi:21744172.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GQET
                Center clone name: CH230-213N13
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 217639 bases at least Q40
                Consensus quality: 220610 bases at least Q30
                Consensus quality: 222513 bases at least Q20
                Estimated insert size: 245420; sum-of-contigs estimation
                Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    11652: contig of 11652 bp in length
            *    11653    11752: gap of unknown length
            *    11753   257149: contig of 245397 bp in length
            *   257150   257249: gap of unknown length
            *   257250   273331: contig of 16082 bp in length.
FEATURES             Location/Qualifiers
     source          1. .273331
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-213N13"
     misc_feature    1. .1233
                     /note="wgs_contig"
     misc_feature    7419. .8672
                     /note="wgs_contig"
     misc_feature    8723. .11652
```

```
                     /note="wgs_contig"
     misc_feature    11753. .14426
                     /note="wgs_contig"
ORIGIN

  Query Match             89.5%;  Score 17;  DB 2;  Length 273331;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db     211057 TTATTGATTCTGTAGGT 211073


RESULT 44
AC106456
LOCUS       AC106456               292913 bp    DNA     linear   HTG 26-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-155P20, *** SEQUENCING IN PROGRESS
            ***, 30 unordered pieces.
ACCESSION   AC106456
VERSION     AC106456.3  GI:23321494
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 292913)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 292913)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 292913)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 26, 2002 this sequence version replaced gi:21735231.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequening reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/).  As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads.  Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GKYF
                Center clone name: CH230-155P20
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 234332 bases at least Q40
                Consensus quality: 242106 bases at least Q30
                Consensus quality: 247333 bases at least Q20
                Estimated insert size: 297300; sum-of-contigs estimation
                Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 30 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     2901: contig of 2901 bp in length
            *    2902     3001: gap of unknown length
            *    3002     6491: contig of 3490 bp in length
            *    6492     6591: gap of unknown length
            *    6592    73160: contig of 66569 bp in length
            *   73161    73260: gap of unknown length
            *   73261    78382: contig of 5122 bp in length
            *   78383    78482: gap of unknown length
            *   78483   176750: contig of 98268 bp in length
            *  176751   176850: gap of unknown length
            *  176851   197686: contig of 20836 bp in length
            *  197687   197786: gap of unknown length
            *  197787   240917: contig of 43131 bp in length
            *  240918   241017: gap of unknown length
            *  241018   244561: contig of 3544 bp in length
            *  244562   244661: gap of unknown length
            *  244662   247987: contig of 3326 bp in length
            *  247988   248087: gap of unknown length
            *  248088   252207: contig of 4120 bp in length
            *  252208   252307: gap of unknown length
```

```
  *   252308   256270: contig of 3963 bp in length
  *   256271   256370: gap of unknown length
  *   256371   259331: contig of 2961 bp in length
  *   259332   259431: gap of unknown length
  *   259432   260877: contig of 1446 bp in length
  *   260878   260977: gap of unknown length
  *   260978   262525: contig of 1548 bp in length
  *   262526   262625: gap of unknown length
  *   262626   263635: contig of 1010 bp in length
  *   263636   263735: gap of unknown length
  *   263736   265033: contig of 1298 bp in length
  *   265034   265133: gap of unknown length
  *   265134   266572: contig of 1439 bp in length
  *   266573   266672: gap of unknown length
  *   266673   267949: contig of 1277 bp in length
  *   267950   268049: gap of unknown length
  *   268050   269166: contig of 1117 bp in length
  *   269167   269266: gap of unknown length
  *   269267   270318: contig of 1052 bp in length
  *   270319   270418: gap of unknown length
  *   270419   272195: contig of 1777 bp in length
  *   272196   272295: gap of unknown length
  *   272296   273360: contig of 1065 bp in length
  *   273361   273460: gap of unknown length
  *   273461   274814: contig of 1354 bp in length
  *   274815   274914: gap of unknown length
  *   274915   276430: contig of 1516 bp in length
  *   276431   276530: gap of unknown length
  *   276531   278517: contig of 1987 bp in length
  *   278518   278617: gap of unknown length
  *   278618   280171: contig of 1554 bp in length
  *   280172   280271: gap of unknown length
  *   280272   282273: contig of 2002 bp in length
  *   282274   282373: gap of unknown length
  *   282374   283669: contig of 1296 bp in length
  *   283670   283769: gap of unknown length
  *   283770   287104: contig of 3335 bp in length
  *   287105   287204: gap of unknown length
  *   287205   292913: contig of 5709 bp in length.
FEATURES             Location/Qualifiers
     source          1. .292913
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-155P20"
     misc_feature    1. .1246
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    1780. .2901
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    3002. .4044
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    3642. .21825
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH365228"
     misc_feature    4700. .6491
                     /note="wgs_contig"
     misc_feature    6592. .7882
                     /note="wgs_contig"
     misc_feature    46254. .46724
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH365230"
     misc_feature    55779. .57065
                     /note="wgs_end_extension

  Query Match             89.5%;  Score 17;  DB 2;  Length 292913;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db       3657 TTATTGATTCTGTAGGT 3673


RESULT 45
AE017014/c
LOCUS       AE017014               300854 bp    DNA     linear   BCT 16-MAY-2003
DEFINITION  Bacillus cereus ATCC 14579 section 17 of 18 of the complete genome.
ACCESSION   AE017014 AE016877
VERSION     AE017014.1  GI:29898525
KEYWORDS    .
SOURCE      Bacillus cereus ATCC 14579
  ORGANISM  Bacillus cereus ATCC 14579
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1  (bases 1 to 300854)
  AUTHORS   Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B.,
            Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N.,
            Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M.,
            Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
            Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
  TITLE     Genome sequence of Bacillus cereus and comparative analysis with
            Bacillus anthracis
  JOURNAL   Nature 423 (6935), 87-91 (2003)
  MEDLINE   22608415
   PUBMED   12721630
REFERENCE   2  (bases 1 to 300854)
  AUTHORS   Candelon,B., Gailloux,K., Ehrlich,D.S. and Sorokin,A.
  TITLE     The number of ribosomal RNA operons in Bacillus cereus
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 300854)
  AUTHORS   Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B.,
            Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N.,
            Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M.,
            Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
            Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de
            Vilvert, Jouy en Josas 78352, France
FEATURES             Location/Qualifiers
     source          1. .300854
                     /organism="Bacillus cereus ATCC 14579"
                     /mol_type="genomic DNA"
                     /strain="ATCC 14579"
                     /db_xref="ATCC:14579"
                     /db_xref="taxon:226900"
     gene            complement(91. .1302)
                     /locus_tag="BC4925"
     CDS             complement(91. .1302)
                     /locus_tag="BC4925"
                     /EC_number="1.6.99.3"
                     /codon_start=1
                     /transl_table=11
                     /product="NADH dehydrogenase"
                     /protein_id="AAP11798.1"
                     /db_xref="GI:29898526"
                     /translation="MVKTPKIVVLGAGYGGMITTVRLQKALSVNEAEITLVNNNSYHY
                     QATWLHESAAGTLQDEKICLDIQDVIDTNKVNFVQDTVVEIKAAEKRIILKDGELEYD
                     YLVIGLGFESETFGITGLKEHAFSIANINATREIREHMEASFSKYATEQRDELVTIVV
                     GGAGFTGIEYVGELANRVPELCKEYNVPREKARIICVEAAPTALPGFDPALVEYAVKQ
                     LEKKGVEFRIGTAIKEATEEGIIVANGDDTELIKSETVVWAAGVRGNGIVEESGFEAM
                     RGRVKVDEYMHAPGYEDVFMVGDAALIINEEINRPYPPTAQIAIQQGYNIAHNLTVLV
                     RGKGEMKKFAFDNKGSVCSLGHDDAMGVVMGKKLTGWKASFMKKVIDNRYLFLLGGPL
                     LVLKKGKLKFF"
     gene            1755. .2750
                     /locus_tag="BC4926"
     CDS             1755. .2750
                     /locus_tag="BC4926"
                     /EC_number="1.6.4.5"
                     /codon_start=1
```

/transl_table=11
/product="Thioredoxin reductase"
/protein_id="AAP11799.1"
/db_xref="GI:29898527"
/translation="MKVAENQKVYDITIIGGGPTGLFTAFYGGMRQASVKIIESLPQL
GGQLSALYPEKYIYDVAGFPKVRAQELVDNLKEQMKKFDPTVCLEEAVDTLEKQADGI
FKLVTNKQTHYSKSVIITAGNGAFQPRRLELEGTAKYEKKNLHYFVDDMNKFAGKRVV
VFGGGDSAVDWTMMLEPIAEKVTIVHRRDKFRAHEHSVENLMNSRAEVSTPYVPVELI
GDDKIEQVVLQHVKTEEKVIIDVDDVIVNYGFVSSLGPIKNWGLDIQKNSILVNSKME
TNIPGIYAAGDICTYEGKVKLIACGFGEAPTAVNNAKAYFDPNAKLQPMHSSSMF"

gene complement(2865. .12986)
/locus_tag="BC4927"

CDS complement(2865. .12986)
/locus_tag="BC4927"
/codon_start=1
/transl_table=11
/product="Cell surface protein"
/protein_id="AAP11800.1"
/db_xref="GI:29898528"
/translation="MKQVLKCFNVVTMITLLLSIILPSPSAFAEINKKSWDELQGNDV
KIKAEQVNAEIYKGEVAKREHVFQYTIENTGATPIQELVLKQNNENEITFLSQSVKVN
GEKLPENKVADFYVEEKDKGGKVLSSNLKIRDLKSHEKMTVLIKASRTQHFNKEYKQK
ISVQKDKVQIGNMSVDVEGIPSEEKVAANADEEADSSKKEVGPSANNTIKKVDEKLKV
SVEDKNNPTSKTELPKQPVTEAEKQPEKEAVKQNVAEVQAAPENASKVQIQTGDKQGF
GDPLYSTITAGNLVQTGNVTLGLTTDHYGRQSLGYKTNKMTVDVDNDDSTFNSSTGAF
PTIPAGSKVKKAYLFWTAAMGTPTYPSKEYRVSDADVREPVKMKMGNKGYAEVSADSI
RKADYLPYISNYSGAGAGYVAYADVTNVIAQQGISQTVTVANVPQIKFEKGSGYYWGN
WNLILVYENHKETVKDMKIWEGLVSQKSTAWTDISVNKINTPKEGAFKAKFSYFSSQG
DPAEKDGYAYDYGEYDFGAGYIKIKNINGKDNDANDSSMTEVQVDGTNEFVTKKYPGY
NPDWTNSFSTDIHTYHLEGPTQVKNDLKEGKMRFRAYAGTGDIYVLNNATFVTEHNAP
NLQVDKKALDSTGKEIKEIKAGEEFTYQIEVKNDTKNNQAPVSNVKGFDKLDDRLEYV
PGSIQYISGSNKGNKTDDASDDEAEFVNNQIDFRVGEGANAKDGGVLKPGESAIITFK
VKVKASVQSDTTVKNIVVVKGQDSTEVNYETADDADITVTTPKEIPGEIEAKKIASNK
TPKLGEEVEYRITFKNKTKDGRLDVLTIEDELPSNLEYVKDSLKAEGAKPEPVELKFE
NGKVIAKYPEIMDMEERSIVFKTKVKETAKIGEEIVNKATVSDKTNPPKNLEEKITPQ
HKAGKIAAKKKATNKKPKLGEEVEYQISFQNTVENGKLDAVTIEDEIPANLEFVQGSE
RAEGAEPNPVELKVENGKVIAKYPEIMDTKERSIAFKVKVKEEAKAGETIVNKAVVSE
PNGQTEHPEEKITPDYKYGKVDVEKSVTNQTPKLGEEVEYRITFYNTVENGKLEKVLV
EDTLPKGVEYVKDSLKAEGIKPEPVELKVEDGKVMAKYPEITDTEKRSIVFKVKVKDS
AKVGEAIVNKAIAKDPKNEPVESKVVITPQSKKGEIAATKVVNNKKPKLGEEIEYRIS
FHNTVENGKLEEVRVEDTIPKGLEYVENSIKAEGEAPGPVELAVENGIVKAKYIDIMD
MKERSIVFKVKVKEEAKVDKEIVNKAIVDDTKNPPIEPEERITPQHKDGIIDSKKTVD
NPSPKLGEEVEYRISFKNTVENGKLEKVIIEDTIPNGLEYVKGSEKAEGDKPAPLKLS
VKDGKVIAEYENITDMKERSIVFKVKVKEEAEVGKEIINKAIVDDMKHPKEPEAKITP
LHKDGKIKAKKSVNNETPKLGEEVEYRISMKNTVKYGKLTDVIIEDTLPEGLEYVENS
LKAEGVGTDSVELKFENGKVMAKYPEIMDTEERSIVFKVKVKEGVTVGKKIINKATID
DSQNKPVNPTAEIIPQYKDGKLEAKKMVNNETPKLGEEVEYRISFKNTVEHGKLTEVK
IEDDLPNGLEYVKDSVKVEGSKPDPVELKFENGKVMAKYPEITDTKERSITFKVKVKG
KVSDSIVNEAIVSDTKHPPETPTAEIIPQHKDGKVKAKKTVNNETPKLGEEVEYRISF
KNTVEDGKLAEVKIEDTLPEGLEYVENSVKAEGSKPDPVELKFENGKVMAKYPEITDT
EERSITFKVKVKDEVKVGKKIVNKAIIDDTKNEPETPTAEITPQHKDGKVEAKKTVNN
ETPKLGEEVEYRISFKNTVENGKLAEVTIKDDLPNGLEYVKDSLRAEGSKPDPVELKV
KDGKVIAKYPEIMDTEERSIVFKAKVKEEFKVGEGIVNKVVVDDTKDPTTSEVTITPE
YKDGKLKAEKFVNNKKPKLGEEVEYRISFKNTVENGKLVEVTVEDEIPAGLEYVENSL
QAEGSKPSPVELKFENGKVMAKYLEIADTKERSIIFKVKVKEEAEIGKEIVNKAIVVD
TKNEPEEPRVEITPQYKDGKIAAQKVANNHKPKLGEEVEYRISFKNTVENGKLVEVNI
KDTLPNGLEYVEGSITAEGSKPKPVELQVKNNKVMAKYPEITDTEERSIVFKAKVKES
VKVGEEIVNEAIIDDTKNLPEEPYVPITPQYKDGKIEARKEVSNHEPKLGEEIEYRII
FNNTVKDGKLAEVKIEDEIPAGLEFVEGSEKAEGEEPKPVELKVENGKVMAKYSEIMD
TKERSIVFKVKVKDSVAIGKDITNKAIIHVDDPNHPVIDPTAKITPQYKDGKIATHKK
VNNHKPKLGEEIEYRISFNNTVKDGKLAEVKIEDEIPSGLEYVKDSLKAEGDKPAPVE
LKEEAGKVSAKYENITDMKERSIIFKVKVKDSVEVDKAIVNKAIVDDTKNPPERPEVE
ITPQHKDGKVKADKKVNKKDPKLGEEVEYRISFKNTVENGKLAEVKIEDQLPTGLEYV
KDSLKSEGNEPAPVELKEEAGKITAKYENIIDTAERSIIFKVKVTEVAKVGKKIVNTA
IVDDDNPKNPPQKPEAIITPEYKDGKLKAEKTVNNPAPKLGEEVEYRITFRNVVEHGK
VAKVKIKDELPTGLEFVEGSERAEGENPKPLYVKVKNGIVLAEYPEITDMKERSIVFK
VKVKEKAKVGEEIVNKAIVEDTINPPEQPNISIQPQYKDGALQAEKTVSNHEPKLGEE
VEYRISFENTVENGKLTEVKVEDEIPAGLEYVQDSIKSDGPEPYPVELKVENGKVTAK
YPEITDTKKRSIIFKVRVQETVQVGKEIVNKAIVDDNNPTPPPVESLIPITPQYKDGK
LEARKEVSNHEPKLGEEIEYRITFNGTVDGGKLVDVKIEDEIPSGLEYVKDSLEAVGD
KPVPTELKVENGKVTVKYPEITDTKERSVIFKVKVKETAKVGGEITNKAIIEVDDPKH
PVIEPTAKIKPEYKDGKLTAAITVSNKEPKLGEEVEYRISFKNTGENGKLDVVTIEDE
IPAGLEYVQGSIRFEGAEPNPIELKMEFGKVIAKYLEITDTKERSIIFNAKVKEAPSK
GITNKAVVDDKINPPLEPTVTILPKTKEDFKIPEDPKEPKEPEVKPEDPKEPKEPEVK
PEDPKEPKEPEVKPEDPKEPKEPEVKPEDPKEPKEPEVKPEDPKEPKEPEVKPEDPKE
PKEPEVKPEDPKEPKEPEVKPEDPKEPKEPEVKPEDPKEPKEPEVKPEDPKEPKEPEV
KPEDPKEPKEPEVKPEDPKEPKEPEVKPEVRLEKLIKEPQVKTERELPKTGAASPWMM
SVGAGISFLVGGVLFVLGRRRKQ"

gene 13424. .13903
/locus_tag="BC4928"

CDS 13424. .13903
/locus_tag="BC4928"
/codon_start=1
/transl_table=11
/product="DNA-binding protein YobU"
/protein_id="AAP11801.1"
/db_xref="GI:29898529"
/translation="MINEPNIMKKEAFQAIGVSVTTTNEREASTKGKISELWNRYFQE
QMMHHIPNKKTKATFAFYSNYESDETGTYTFTIGTTVSSLEDVPENMTTLTIPAATYA
VFTTRKGPVAEVVCEAWEYIWQWSKENKRAFTTDFELYDERALDPNHAQLDIYIALA"

gene 14121. .15368
/locus_tag="BC4929"

CDS 14121. .15368
/locus_tag="BC4929"
/codon_start=1
/transl_table=11
/product="Macrolide-efflux protein"
/protein_id="AAP11802.1"
/db_xref="GI:29898530"
/translation="MEELQQNKTALEGSGKPLLKNTNFLFLWAATLFSSFALAFFTFS
QTWYIAKTLNLEASLGVVFVALSVPRLIFMIVGGAVADKFPKKNIMFYSNIIRAILVA
TILTWFIVGDITLYTFALFALFFGLADAFFWSADGSILPELVEKSRLTQANSLTQMTN
QASVILGPVLGGILIKFTNYETIFSITILLLIVAAILVQKIQFTMPEQKDTDKGMFTS
IKEGILYVKESPFLSTFLICSAFLNLFLIGPMQVGFPLFVKNVLHGDSLQFSYLEASV
GGGMAIGAVIVGLKNINRRRGLFCIIMMLLSGVFFLSINFSTVLWQALLAGMFYGITI
AMAIVPLMAMIQSTVKEEMMGRVMSLLMLSSMGFIPLSYAFTSIALAIGIPIVTVMKS
GAIAVIVFVLFVAIRVPVVRKFD"

gene complement(15379. .16266)
/locus_tag="BC4930"

CDS complement(15379. .16266)
/locus_tag="BC4930"
/codon_start=1
/transl_table=11
/product="Transcriptional regulator, DeoR family"
/protein_id="AAP11803.1"
/db_xref="GI:29898531"
/translation="MGSLYNLMKPYSEFQSKEEFNTYQKQVLKHYRFQLNKTDATIIH
FLGKYAVNEKQKTVGVACPLMETIATNVGKSIRTVRRSIAKLEELGIIKRVATKERHR
RGGYSANLYVFLISAIDRMDDRMKMSACESEDYTAGCSGKEQKSERETILSKNIPQIK
EKRKITYELDETYCRHDIPKPFIYALLPMTSNPKKINIFWSKVELAYKKSGLLEQGVL
LEQILADEEVYGNFIWRVKSVVRAYKYGEIRKNVKALLYSTVRDLFLEVGLEWGAAVR
RSKGIPLFDPFKKEPCVNA"

gene 16347. .16808
/locus_tag="BC4931"

CDS 16347. .16808
/locus_tag="BC4931"
/codon_start=1
/transl_table=11
/product="hypothetical Cytosolic Protein"
/protein_id="AAP11804.1"
/db_xref="GI:29898532"
/translation="MIVISACLGGIACRYDGNDNLVSKIEELLQKEDTVLICPEVLGG
LPTPRPSAEIIGGNGDDVLDGKAKVMTKDGEDVTEAFVSGAYKALEQIKNLHPEYIIL
KERSPSCGSSTIYTGEFNGNKQTGFGVTTALFKRHGFKVISEEEFENKKRN"

gene complement(16839. .17192)
/locus_tag="BC4932"

Query Match 89.5%; Score 17; DB 1; Length 300854;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAG 17
||||||||||||||||||
Db 169021 AATTATTGATTCTGTAG 169005

Search completed: April 26, 2005, 16:01:54
Job time : 1826 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 26, 2005, 16:02:03 ; Search time 500 Seconds
(without alignments)
230.700 Million cell updates/sec

Title: US-10-619-906-5
Perfect score: 19
Sequence: 1 aattattgattctgtaggt 19

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5633728 seqs, 3035525691 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		19	100.0	19	17	US-10-619-906-5	Sequence 5, Appli
c 2		19	100.0	458	9	US-09-920-455-179	Sequence 179, App
c 3		19	100.0	589	11	US-09-969-034-2908	Sequence 2908, Ap
c 4		19	100.0	1413	16	US-10-337-632-3	Sequence 3, Appli
c 5		19	100.0	1413	17	US-10-295-027-323	Sequence 323, App
c 6		19	100.0	1413	17	US-10-295-027-1026	Sequence 1026, Ap
c 7		19	100.0	1413	17	US-10-295-027-1027	Sequence 1027, Ap
c 8		19	100.0	1653	18	US-10-384-339C-101	Sequence 101, App
c 9		19	100.0	1778	9	US-09-954-456-53	Sequence 53, Appl
c 10		19	100.0	1778	9	US-09-954-456-1127	Sequence 1127, Ap
c 11		19	100.0	1778	9	US-09-954-456-1799	Sequence 1799, Ap
c 12		19	100.0	1778	9	US-09-880-107-2179	Sequence 2179, Ap
c 13		19	100.0	1778	9	US-09-920-455-218	Sequence 218, App
c 14		19	100.0	1778	10	US-09-873-367C-439	Sequence 439, App
c 15		19	100.0	1778	10	US-09-873-367C-1001	Sequence 1001, Ap
c 16		19	100.0	1778	14	US-10-171-311-134	Sequence 134, App
c 17		19	100.0	1778	15	US-10-301-822-122	Sequence 122, App
c 18		19	100.0	1778	16	US-10-337-632-1	Sequence 1, Appli
c 19		19	100.0	1778	16	US-10-131-985-40	Sequence 40, Appl
c 20		19	100.0	1778	17	US-10-172-118-769	Sequence 769, App
c 21		19	100.0	1778	17	US-10-409-643-14	Sequence 14, Appl
c 22		19	100.0	1778	17	US-10-240-425-1283	Sequence 1283, Ap
c 23		19	100.0	1778	17	US-10-342-887-769	Sequence 769, App
c 24		19	100.0	1778	17	US-10-188-832-124	Sequence 124, App
c 25		19	100.0	1778	17	US-10-619-906-1	Sequence 1, Appli
c 26		19	100.0	1778	18	US-10-734-564-10	Sequence 10, Appl
c 27		19	100.0	1778	18	US-10-751-736-12	Sequence 12, Appl
c 28		19	100.0	1778	19	US-10-901-417-40	Sequence 40, Appl
c 29		19	100.0	1778	19	US-10-843-641A-439	Sequence 439, App
c 30		19	100.0	1778	19	US-10-843-641A-1001	Sequence 1001, Ap
c 31		19	100.0	1778	19	US-10-843-641A-3080	Sequence 3080, Ap
c 32		19	100.0	1778	19	US-10-843-641A-4154	Sequence 4154, Ap
c 33		19	100.0	1778	19	US-10-843-641A-4826	Sequence 4826, Ap
c 34		19	100.0	1873	18	US-10-723-860-6186	Sequence 6186, Ap
c 35		19	100.0	1874	9	US-09-925-301-375	Sequence 375, App
c 36		19	100.0	1988	10	US-09-971-392-36	Sequence 36, Appl
c 37		19	100.0	2870	18	US-10-335-053-279	Sequence 279, App
c 38		17	89.5	47	17	US-10-349-143-2352	Sequence 2352, Ap
39		16	84.2	65	16	US-10-032-585-3874	Sequence 3874, Ap
40		16	84.2	1116	18	US-10-437-963-37801	Sequence 37801, A
c 41		16	84.2	1554	16	US-10-032-585-6874	Sequence 6874, Ap
42		16	84.2	38653	16	US-10-220-510-11	Sequence 11, Appl
c 43		15	78.9	25	19	US-10-719-900-37891	Sequence 37891, A
44		15	78.9	25	19	US-10-719-900-701443	Sequence 701443,
45		15	78.9	274	16	US-10-029-386-23884	Sequence 23884, A

ALIGNMENTS

RESULT 1
US-10-619-906-5
; Sequence 5, Application US/10619906
; Publication No. US20040087533A1
; GENERAL INFORMATION:
; APPLICANT: Index Pharmaceuticals
; TITLE OF INVENTION: New Compound
; FILE REFERENCE: 50299
; CURRENT APPLICATION NUMBER: US/10/619,906
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(19)
; OTHER INFORMATION: SEQ ID NO. 5, antisense oligonucleotide
US-10-619-906-5

Query Match 100.0%; Score 19; DB 17; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 1 AATTATTGATTCTGTAGGT 19

RESULT 2
US-09-920-455-179/c
; Sequence 179, Application US/09920455

; Patent No. US20020168647A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF HEAD AND NECK CANCER
; FILE REFERENCE: 210121.540
; CURRENT APPLICATION NUMBER: US/09/920,455
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 179
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-455-179

Query Match 100.0%; Score 19; DB 9; Length 458;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 348 AATTATTGATTCTGTAGGT 330

RESULT 3
US-09-969-034-2908/c
; Sequence 2908, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2908
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 456, 463, 481, 541, 545, 572, 576, 586, 589
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-2908

Query Match 100.0%; Score 19; DB 11; Length 589;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 419 AATTATTGATTCTGTAGGT 401

RESULT 4
US-10-337-632-3/c
; Sequence 3, Application US/10337632
; Publication No. US20030157110A1
; GENERAL INFORMATION:
; APPLICANT: An, Frank
; APPLICANT: Chen, Hong
; TITLE OF INVENTION: METHODS FOR THE TREATMENT OF METABOLIC
; TITLE OF INVENTION: DISORDERS, INCLUDING OBESITY AND DIABETES
; FILE REFERENCE: MPI01-253P1RM
; CURRENT APPLICATION NUMBER: US/10/337,632
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: 60/346,354
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-337-632-3

Query Match 100.0%; Score 19; DB 16; Length 1413;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 361 AATTATTGATTCTGTAGGT 343

RESULT 5
US-10-295-027-323/c
; Sequence 323, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 323
; LENGTH: 1413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-323

```
  Query Match             100.0%;  Score 19;  DB 17;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 6
US-10-295-027-1026/c
; Sequence 1026, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1026
;   LENGTH: 1413
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-295-027-1026

  Query Match             100.0%;  Score 19;  DB 17;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 7
US-10-295-027-1027/c
; Sequence 1027, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1027
;   LENGTH: 1413
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-295-027-1027

  Query Match             100.0%;  Score 19;  DB 17;  Length 1413;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 8
US-10-384-339C-101/c
; Sequence 101, Application US/10384339C
; Publication No. US20040175703A1
; GENERAL INFORMATION:
;  APPLICANT: Kreutzer, Roland
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING EXPRESSION OF A TARGET GE
;  FILE REFERENCE: 20200/2002
;  CURRENT APPLICATION NUMBER: US/10/384,339C
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/EP02/00152
;  PRIOR FILING DATE: 2002-01-09
;  PRIOR APPLICATION NUMBER: DE 10100586.5
;  PRIOR FILING DATE: 2001-01-09
;  PRIOR APPLICATION NUMBER: DE 10155280.7
;  PRIOR FILING DATE: 2001-10-26
;  PRIOR APPLICATION NUMBER: DE 10158411.3
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: DE 10160151.4
;  PRIOR FILING DATE: 2001-12-07
```

```
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
;   LENGTH: 1653
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (626)..(925)
;   OTHER INFORMATION: n=A, T, G, C or gap
;  PUBLICATION INFORMATION:
;   TITLE: MMP12
;   PATENT DOCUMENT NUMBER: XM006272
US-10-384-339C-101

  Query Match             100.0%;  Score 19;  DB 18;  Length 1653;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        361 AATTATTGATTCTGTAGGT 343


RESULT 9
US-09-954-456-53/c
; Sequence 53, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
;   LENGTH: 1778
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-53

  Query Match             100.0%;  Score 19;  DB 9;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 10
US-09-954-456-1127/c
; Sequence 1127, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1127
;   LENGTH: 1778
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-1127

  Query Match             100.0%;  Score 19;  DB 9;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              |||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 11
US-09-954-456-1799/c
; Sequence 1799, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1799
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1799

Query Match 100.0%; Score 19; DB 9; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 12
US-09-880-107-2179/c
; Sequence 2179, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2179
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L23808
US-09-880-107-2179

Query Match 100.0%; Score 19; DB 9; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 13
US-09-920-455-218/c
; Sequence 218, Application US/09920455
; Patent No. US20020168647A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF HEAD AND NECK CANCER
; FILE REFERENCE: 210121.540
; CURRENT APPLICATION NUMBER: US/09/920,455
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 218
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-455-218

Query Match 100.0%; Score 19; DB 9; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 14
US-09-873-367C-439/c
; Sequence 439, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 439
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-439

Query Match 100.0%; Score 19; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 15
US-09-873-367C-1001/c
; Sequence 1001, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29

; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1001
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-1001

Query Match 100.0%; Score 19; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 16
US-10-171-311-134/c
; Sequence 134, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-134

Query Match 100.0%; Score 19; DB 14; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 17
US-10-301-822-122/c
; Sequence 122, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)...(1425)
US-10-301-822-122

Query Match 100.0%; Score 19; DB 15; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 18
US-10-337-632-1/c
; Sequence 1, Application US/10337632
; Publication No. US20030157110A1
; GENERAL INFORMATION:
; APPLICANT: An, Frank
; APPLICANT: Chen, Hong
; TITLE OF INVENTION: METHODS FOR THE TREATMENT OF METABOLIC
; TITLE OF INVENTION: DISORDERS, INCLUDING OBESITY AND DIABETES
; FILE REFERENCE: MPI01-253P1RM
; CURRENT APPLICATION NUMBER: US/10/337,632
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: 60/346,354
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)...(1425)
US-10-337-632-1

Query Match 100.0%; Score 19; DB 16; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 19
US-10-131-985-40/c

; Sequence 40, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-985-40

Query Match 100.0%; Score 19; DB 16; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 20
US-10-172-118-769/c
; Sequence 769, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 769
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_002426
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-769

Query Match 100.0%; Score 19; DB 17; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 21
US-10-409-643-14/c
; Sequence 14, Application US/10409643
; Publication No. US20030235577A1
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven
; APPLICANT: Hartzell, William
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PREVENTING AND TREATING MICROBIAL INF
; FILE REFERENCE: B00801.70281.US
; CURRENT APPLICATION NUMBER: US/10/409,643
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: US 60/370,649
; PRIOR FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-409-643-14

Query Match 100.0%; Score 19; DB 17; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 22
US-10-240-425-1283/c
; Sequence 1283, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1283
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 L23808
US-10-240-425-1283

Query Match 100.0%; Score 19; DB 17; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 23
US-10-342-887-769/c
; Sequence 769, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 769
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-769

Query Match 100.0%; Score 19; DB 17; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 24
US-10-188-832-124/c
; Sequence 124, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 124
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-124

Query Match 100.0%; Score 19; DB 17; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 25
US-10-619-906-1/c
; Sequence 1, Application US/10619906
; Publication No. US20040087533A1
; GENERAL INFORMATION:
; APPLICANT: Index Pharmaceuticals
; TITLE OF INVENTION: New Compound
; FILE REFERENCE: 50299
; CURRENT APPLICATION NUMBER: US/10/619,906
; CURRENT FILING DATE: 2003-07-16
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1778)
; OTHER INFORMATION: SEQ ID NO.1; cDNA human MMP-12, Genebank acc. no. NM-002426
US-10-619-906-1

Query Match 100.0%; Score 19; DB 17; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 26
US-10-734-564-10/c
; Sequence 10, Application US/10734564
; Publication No. US20040157278A1
; GENERAL INFORMATION:
; APPLICANT: Christopher C Burgess et al
; TITLE OF INVENTION: Detection Methods Using TIMP1
; FILE REFERENCE: 1657/2012
; CURRENT APPLICATION NUMBER: US/10/734,564
; CURRENT FILING DATE: 2003-12-12
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-734-564-10

Query Match 100.0%; Score 19; DB 18; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 27
US-10-751-736-12/c
; Sequence 12, Application US/10751736
; Publication No. US20040265230A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM100927 (031896-002000)
; CURRENT APPLICATION NUMBER: US/10/751,736

; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; PRIOR FILING DATE: 2003-01-06
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-751-736-12

Query Match 100.0%; Score 19; DB 18; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 28
US-10-901-417-40/c
; Sequence 40, Application US/10901417
; Publication No. US20050026836A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/901,417
; CURRENT FILING DATE: 2004-07-28
; PRIOR APPLICATION NUMBER: US/10/131,985
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-901-417-40

Query Match 100.0%; Score 19; DB 19; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 29
US-10-843-641A-439/c
; Sequence 439, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 439
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-439

Query Match 100.0%; Score 19; DB 19; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATTATTGATTCTGTAGGT 19
||||||||||||||||||||
Db 373 AATTATTGATTCTGTAGGT 355

RESULT 30
US-10-843-641A-1001/c
; Sequence 1001, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1001
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-1001

```
  Query Match             100.0%;  Score 19;  DB 19;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 31
US-10-843-641A-3080/c
; Sequence 3080, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3080
;   LENGTH: 1778
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-3080

  Query Match             100.0%;  Score 19;  DB 19;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 32
US-10-843-641A-4154/c
; Sequence 4154, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4154
;   LENGTH: 1778
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-4154

  Query Match             100.0%;  Score 19;  DB 19;  Length 1778;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355


RESULT 33
US-10-843-641A-4826/c
; Sequence 4826, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4826
;   LENGTH: 1778
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-10-843-641A-4826

Query Match 100.0%; Score 19; DB 19; Length 1778;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        373 AATTATTGATTCTGTAGGT 355
```

RESULT 34
US-10-723-860-6186/c
; Sequence 6186, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6186
; LENGTH: 1873
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-6186

Query Match 100.0%; Score 19; DB 18; Length 1873;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        421 AATTATTGATTCTGTAGGT 403
```

RESULT 35
US-09-925-301-375/c
; Sequence 375, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 375
; LENGTH: 1874
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-375

Query Match 100.0%; Score 19; DB 9; Length 1874;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        423 AATTATTGATTCTGTAGGT 405
```

RESULT 36
US-09-971-392-36/c
; Sequence 36, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 1988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Template ID: 200569.1
; NAME/KEY: unsure
; LOCATION: 1886
; OTHER INFORMATION: a, t, c, g, or other
US-09-971-392-36

Query Match 100.0%; Score 19; DB 10; Length 1988;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db        396 AATTATTGATTCTGTAGGT 378
```

RESULT 37
US-10-335-053-279/c
; Sequence 279, Application US/10335053
; Publication No. US20040241653A1
; GENERAL INFORMATION:
; APPLICANT: Quark Biotech, Inc.
; TITLE OF INVENTION: Methods for identifying marker genes for cancer
; FILE REFERENCE: 68733-A; 070/US1
; CURRENT APPLICATION NUMBER: US/10/335,053
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: 60/345,317
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 279
; LENGTH: 2870
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-335-053-279

Query Match 100.0%; Score 19; DB 18; Length 2870;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AATTATTGATTCTGTAGGT 19
              ||||||||||||||||||||
Db       1465 AATTATTGATTCTGTAGGT 1447
```

RESULT 38
US-10-349-143-2352/c
; Sequence 2352, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel

```
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/10/349,143
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: US/09/422,978
;  PRIOR FILING DATE: 1999-10-20
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2352
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-10570-107 : polymorphic base G or A
US-10-349-143-2352

  Query Match             89.5%;  Score 17;  DB 17;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 5.5;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TTATTGATTCTGTAGGT 19
              |||||||||||||||||
Db         21 TTATTGATTCTGTAGGT 5


RESULT 39
US-10-032-585-3874
; Sequence 3874, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT:  Bo, Jiang
;  APPLICANT:  Charles, Boone
;  APPLICANT:  Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3874
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-032-585-3874

  Query Match             84.2%;  Score 16;  DB 16;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db         39 TATTGATTCTGTAGGT 54


RESULT 40
US-10-437-963-37801
; Sequence 37801, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 37801
;   LENGTH: 1116
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_41497C.1
US-10-437-963-37801

  Query Match             84.2%;  Score 16;  DB 18;  Length 1116;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db        964 TATTGATTCTGTAGGT 979


RESULT 41
US-10-032-585-6874/c
; Sequence 6874, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT:  Bo, Jiang
;  APPLICANT:  Charles, Boone
;  APPLICANT:  Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6874
;   LENGTH: 1554
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-032-585-6874

  Query Match             84.2%;  Score 16;  DB 16;  Length 1554;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TATTGATTCTGTAGGT 19
              ||||||||||||||||
Db         27 TATTGATTCTGTAGGT 12


RESULT 42
US-10-220-510-11
; Sequence 11, Application US/10220510
; Publication No. US20030190637A1
; GENERAL INFORMATION:
;  APPLICANT: Hovnanian, Alain
;  APPLICANT:  Chavanas, Stephane
;  APPLICANT:  Cookson, William
;  APPLICANT:  Moffat, Miriam
;  APPLICANT:  Walley, Andrew
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR NETHERTON'S DISEASE
;  FILE REFERENCE: I00317.70008.US
;  CURRENT APPLICATION NUMBER: US/10/220,510
;  CURRENT FILING DATE:  2003-03-04
;  PRIOR APPLICATION NUMBER: GB 0005098.9
;  PRIOR FILING DATE: 2000-03-02
```

; PRIOR APPLICATION NUMBER: GB 0005229.0
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 38653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1846)..(1846)
; OTHER INFORMATION: n = a, c, g or t/u
US-10-220-510-11

Query Match 84.2%; Score 16; DB 16; Length 38653;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TATTGATTCTGTAGGT 19
||||||||||||||||
Db 31420 TATTGATTCTGTAGGT 31435

RESULT 43
US-10-719-900-37891/c
; Sequence 37891, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 37891
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-37891

Query Match 78.9%; Score 15; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTATTGATTCTGTAG 17
|||||||||||||||
Db 21 TTATTGATTCTGTAG 7

RESULT 44
US-10-719-900-701443
; Sequence 701443, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 701443
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-701443

Query Match 78.9%; Score 15; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGTA 16
|||||||||||||||
Db 4 ATTATTGATTCTGTA 18

RESULT 45
US-10-029-386-23884
; Sequence 23884, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23884
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004237.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: SWISSPROT HIT: Q9Y496, EVALUE 6.00e-49
; OTHER INFORMATION: NT HIT: gi14781205, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BI561517.1, EVALUE 0.00e+00
US-10-029-386-23884

Query Match 78.9%; Score 15; DB 16; Length 274;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ATTATTGATTCTGTA 16
|||||||||||||||
Db 244 ATTATTGATTCTGTA 258

Search completed: April 26, 2005, 17:05:13
Job time : 503 secs